University College London

Department of Medical Physics & Biomedical Engineering

Proton and Advanced RadioTherapy Group

PhD Thesis
submitted for the degree of Doctor of Philosophy from University College London

Toward adaptive radiotherapy

Catarina Isabel Correia Veloso da Veiga

Supervisors: Gary Royle – Proton and Advanced Radiotherapy Group, UCL
Jamie McClelland – Centre for Medical Image Computing, UCL
Kate Ricketts – Division of Surgery and Interventional Sciences, UCL

2016

Declaration

I, Catarina Isabel Correia Veloso da Veiga, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm this has been indicated in the thesis.

Signed .....

Abstract

Intensity Modulated Radiotherapy (IMRT) and proton therapy are the state-of-art external radiotherapy modalities. To make the most of such precise delivery, accurate knowledge of the patient anatomy and biology during treatment is necessary, as unaccounted variations can compromise the outcome of the treatment. Treatment modification to account for deviations from the planning stage is a framework known as adaptive radiotherapy (ART).

To fully utilise the information extracted from different modalities and/or at different time-points it is required to accurately align the imaging data. In this work the feasibility of cone-beam computed tomography (CBCT) and deformable image registration (DIR) for ART was evaluated in the context of head and neck (HN) and lung malignancies, and for IMRT and proton therapy applications. This included the geometric validation of deformations for multiple DIR algorithms, estimating the uncertainty in dose recalculation of a CBCT-based deformed CT (dCT), and the uncertainty in dose summation resulting from the properties of the underlying deformations. The dCT method was shown to be a good interim solution to repeat CT and a superior alternative to simpler direct usage of CBCT for dose calculation; proton therapy treatments were more sensitive to registration errors than IMRT. The ability to co-register multimodal and multitemporal data of the HN was also explored; the results found were promising and the limitations of current algorithms and data acquisition protocols were identified.

The use of novel artificial cancer masses as a novel platform for the study of imaging during radiotherapy was explored in this study. The artificial cancer mass model was extended to generate magnetic resonance imaging (MRI)-friendly samples. The tumoroids were imageable in standard T_1 and T_2 MRI acquisitions, and the relaxometric properties were measured. The main limitation of the current tumour model was the poor reproducibility and controllability of the properties of the samples.

Acknowledgements

> That sad look shake it off, the road has been too long. We're all just passengers, in time and space, an end we never chase.
>
> Tom Barman

I would sincerely like to thank my supervisors, without whom this work would not have been possible. First, to Professor Gary Royle for giving me this unique opportunity along with his constant support, guidance, time and great sense of humour over the course of these long four years. To Dr Jamie McClelland, for all those long coffee breaks discussing papers and code, for all his patience when something went wrong, for all the kindness and trust; thank you for being my rock during the difficult moments of this PhD. And finally, to Dr Kate Ricketts, who always knew how to lift my spirits with her enthusiasm, support and encouragement.

During the duration of my PhD, I was extremely lucky to collaborate with fantastic people from different research groups and backgrounds. From the Centre for Medical Image Computing, I would like to acknowledge Marc Modat, Pankaj Daga, Gergely Zombori, Matt Clarkson, Sebastien Ourselin, Dave Hawkes, and Marcel van Herk for the support received with image registration in general, NifTK in particular and overall direction of this project. From the Departments of Radiotherapy and Radiotherapy Physics at University College London Hospital, I am indebted to Derek D'Souza, Ivan Rosenberg and Richard Amos for always so kindly motivating, encouraging, and guiding me in this journey; and to Rachel Bodey, Syed Moinuddin, Paul Doolan, Jailan Alshaiki, Phil Davies, Chris Stacey, Maria Kilkenny, Dr Dhanasekaran Kittappa, Dr Swee-Ling Wong, and Dr Ruheena Mendes for always finding the time and patience to help me with all the clinical needs of this project. From the Centre for Medical Imaging, I would like to show my appreciation for Dr Shonit Punwani and Heather Fitzke for the helpful discussions on the usability of MRI in radiotherapy and access to clinical trial data. From the Division of Surgery, I would like to thank Dr Marilena Loizidou, Tarig Magdeldin, Tong Long, Victor Lopez-Davila, and Bala Ramesh for all the patience required to guide a physicist through the complex world of biology. I am also very grateful to Bernard Siow, for all the long

hours spent using the pre-clinical MRI scanner at the Centre for Advanced Biomedical Imaging and his great sense of humour through it all. My thanks to Denzil Booth, Robert Moss and George Randall for their amazing workshop skills, Dan O'Flynn for data acquisition on a benchtop CT scanner, and Reem Al-Samarraie for support in locating the materials of the box for transporting the tumoroids. Finally, my appreciation to Amber Cuming (Beekley Corporation) and Samuel Naslund (Naslund Medical AB) for kindly providing samples of CT/MR markers.

From my time at the University of Pennsylvania, I would first like to thank Dr Kevin Teo for receiving me so warmly in Philadelphia and at the Roberts Proton Therapy Center; I always felt very welcome and a valuable member of the team. A very special thanks goes to Guillaume Janssens for his 24/7 support, professionalism and sympathy; I was really lucky to always have the best collaborators. To Dr Ching-Ling Teng for her enthusiasm in the final phase of my stay in Philadelphia; without your excellent writing skills, full-time availability, girly lunch breaks and great attitude I do not believe this project could have been as successful. I am also grateful to Thomas Baudier and Lucian Hotoiu for their kindness and technical contributions to the project, and to Lingshu Yin, Sebastien Brousmiche, Liyong Lin, James Metz, Timothy Solberg, Zelig Tochner, and James McDonough for their contributions to the CBCT and adaptive lung proton therapy project. To David Weiss, Evan Meekins and Marcus Fager, for adding *sweet, pickle* and *meat* to my *saltyness*. To my favourite *guapitas* for cheering me up when I needed it the most. I was really my happiest during my time at UPenn, and I am so grateful to have been given the opportunity to do this project. Looking back I realise how much I grew as a woman and a scientist in the United States. Philly, you changed me for better and for good.

To everyone in the Department of Medical Physics & Biomedical Engineering at UCL, particularly, Alessandro Proverbio, Anna Zamir, Christiana Christodolou, Dan O'Flynn, Edgar Gelover Reyes, Emma Biondetti, Esther Bär, Vanessa La Rosa, George Randall, Ireneos Drakos, and Paul Burke. Thank you for all the good times, either at the office, common room, pub or outdoors! And coffee, thank you for all the coffee!

A very emotional thanks goes to my important friends, my lifetime friends: Abigail Moreira, Ana Luísa Castro Lopes, Ana Mónica Lourenço, Bruno Gomes, Consuelo Guardiola, Diana Barros, João Koch, João Nuno Mota, João Tavares, Juliana Narciso, Lucia Tejo, Mariana Dantas, Nuno Alcobia, Raphael Lopes, Raquel Koch, Raquel Leão, Reem Al-Samarraie, Sara Campos, Sofia C. Ribeiro, and Sofia J. Ribeiro. Distance can take quite a toll on relationships; nevertheless, you were always there for me. Knowing you always waited for my return, making me feel as if I had never left, meant more than words can convey. I am truly blessed to have such amasing people in my life. I would also like to express my genuine appreciation to Paulo Castro and Stavros Vorrias; while our paths may have split somewhere along this long road, your companionship was really important to me throughout our happy times.

My most special thanks goes to my family, who fully supported me through all the good and bad moments, and always cheered for my success even if it was so hard to be apart: my parents Carlos and Isabel; my brother Nuno; my grandparents, Eurico, Maria Augusta and Ulisses; my amasing and favourite aunts, Patricia, Nocas and Sissi; and my little cousin, Eva. I know the moments we lost will never be recovered, but allow me to be selfish this one time and convince myself that we still have all the time in the world to recapture them all.

Finally, I would like to gratefully acknowledge the financial support received from Fundação para a Ciência e a Tecnologia (FCT) grant SFRH/BD/76169/2011, co-financed by ESF, POPH/QREN and EU; IOP, IPEM and UCL Graduate School for funding conference trips.

Dedicated to my grandparents, Eurico and Maria Augusta.

Contents

List of Figures

List of Tables

Glossary

1D unidimensional.

2D bidimensional.

3D tridimensional.

4D four-dimensional.

AAA analytical isotropic algorithm.

ACM artificial cancer mass.

ADC apparent diffusion coefficient.

AP anterior-posterior.

ART adaptive radiation therapy.

BE bending energy penalty term.

BOLD blood oxygen level dependent.

CABI Centre for Advanced Biomedical Imaging.

CBCT cone-beam CT.

CMIC Centre of Medical Image Computing.

CoM center-of-mass (centroid) position error.

CP control point.

CPS control point spacing.

CT computed tomography.

CTV clinical tumour volume.

dCBCT deformed CBCT.

DCE dynamic contrast enhancement.

dCT deformed CT.

DD dose-difference.

DICOM Digital imaging in communications in medicine.

DIR deformable image registration.

DMEM Dulbecco's Modified Eagle Medium.

DNA deoxyribonucleic acid.

DSC dice similarity coefficient.

DT distance transformation.

DTA distance to agreement.

DTD distance to dose-difference.

DVF deformation vector field.

DVH dose-volume histogram.

DW diffusion-weighted.

ECM extracellular matrix.

EDTA ethylenediaminetetraacetic acid.

FBS foetal bovine serum.

FFD free-form deformation.

FID free induction decay.

FLASH fast low angle shot.

FN false negatives.

FoV field-of-view.

FP false positives.

GE gradient echo.

GPU graphics processing unit.

GTV gross tumour volume.

HE harmonic energy.

HN head and neck.

HU Hounsfield unit.

HUP Hospital of the University Pennsylvania.

IBA Ion Beam Applications, SA.

IC inverse-consistency penalty term.

ICE inverse-consistency error.

iCTV internal clinical tumour volume.

IGRT image-guided radiation therapy.

IMPT intensity modulated proton therapy.

IMRT intensity modulated radiation therapy.

IR inversion-recovery.

IV irradiated volume.

JL logarithm of the Jacobian determinant penalty term.

LINAC linear accelerator.

LNCC localised normalised cross correlation.

LSCM left sternocleidomastoid muscle.

MEM minimum essential medium.

MFO multiple-field optimisation.

MIP maximum intensity projection.

MR magnetic resonance.

MRI magnetic resonance imaging.

MRS magnetic resonance spectroscopy.

MSME multi-slice multi-echo.

NCC normalised cross correlation.

NMI normalised mutual information.

OAR organ-at-risk.

OI overlap index.

P/S penicillin/streptomycin.

PA posterior-anterior.

PBS phosphate buffed saline.

pCT planning computed tomography.

pD prescribed dose.

PET positron emission tomography.

PLA polylactic acid.

PSPT passive scattering proton therapy.

PTV planning target volume.

RARE rapid acquisition refocused echoes.

rCBCT regular CBCT.

rCT replan/rescan CT.

RED relative electron density.

RF radio-frequency.

RL right-left.

RMS root mean square.

ROI region of interest.

RSCM right sternocleidomastoid muscle.

RTK Reconstruction Toolbox.

sCBCT simulated CBCT.

SE spin-echo.

SFUD single-field uniform dose.

SI superior-inferior.

SOBP spread-out Bragg peak.

SPECT single-photon emission computed tomography.

SSD sum of the squared differences.

TE echo time.

TI inversion time.

TPS treatment planning system.

TR repetition time.

TRE target registration error.

TV treated volume.

UCL University College London.

UCLH University College London Hospital.

UK United Kingdom.

USA United States of America.

VOI volume of interest.

WET water equivalent thickness.

WHO World Health Organization.

Outputs

Peer-reviewed journal papers

(In preparation) N. Hobson, X. Weng-Jiang, **C. Veiga**, B. Siow, M. Ashford, N. T. K. Thanh, A. Schätzlein, and I. Uchegbu, "Design and synthesis of self-assembling polymeric iron oxide nanoparticle theranostics for applications in cancer diagnostics and cancer therapy," ACS Nano (2016).

(In preparation) X. Weng-Jiang, N. Hobson, **C. Veiga**, B. Siow, N. T. K. Thanh, A. Schätzlein, and I. Uchegbu, "Aqueous in-flow synthesis of superparamagnetic iron oxide nanoparticles for dual T_1/T_2-weighted magnetic resonance imaging," ACS Nano (2016).

(In preparation) **C. Veiga**, G. Janssens, T. Baudier, L. Hotoiu, S. Brousmiche, J. R. McClelland, C.-L. Teng, L. Yin, G. Royle, and B.-K. K. Teo, "The accuracy of CBCT and deformable registration for adaptive lung proton therapy" (2016).

C. Veiga, G. Janssens, C.-L. Teng, T. Baudier, L. Hotoiu, J. R. McClelland, G. Royle, L. Lin, L. Yin, J. Metz, T. D. Solberg, Z. Tochner, C. B. Simone II, J. McDonough, and B.-K. K. Teo, "First clinical investigation of CBCT and deformable registration for adaptive proton therapy of lung cancer," Int. J. Radiat. Oncol. Biol. Phys. 95(1) 549-559 (2016).

C. Veiga, J. Alshaikhi, R. Amos, A. M. Lourenço, M. Modat, S. Ourselin, G. Royle, and J. R. McClelland, "CBCT and deformable registration based "dose of the day" calculations for adaptive proton therapy," Int. J. Particle Ther. 2(2) 404-414 (2015).

A. K. Hoang Duc, G. Eminowicz, J. McClelland, M. Modat, M. J. Cardoso, A. F. Mendelson, **C. Veiga**, T. Kadir, D. D'Souza, and S. Ourselin, "Validation of clinical acceptability of an atlas-based segmentation algorithm for the delineation of organs at risk in head and neck cancer," Med. Phys. 42(9) 5027-5034 (2015).

C. Veiga, A. Lourenço, S. Moinuddin, M. van Herk, M. Modat, S. Ourselin, D. D'Souza, G. Royle, and J. R. McClelland, "Toward adaptive radiotherapy for head and neck patients: uncertainties in dose warping due to the choice of deformable registration algorithm," Med. Phys. 42(2) 760-769 (2015).

C. Veiga, J. McClelland, S. Moinuddin, A. Lourenço, K. Ricketts, J. Annkah, M. Modat, S. Ourselin, D. D'Souza, and G. Royle, "Toward adaptive radiotherapy for head and neck patients: feasibility study on using CT-to-CBCT deformable registration for "dose of the day" calculations," Med. Phys. 41 031703 (2014).

Peer-reviewed conference papers

C. Veiga, R. Mendes, D. Kittapa, S.-L. Wong, R. Bodey, M. Modat, S. Ourselin, G. Royle, and J. McClelland, "Optimization of Multimodal and Multitemporal Deformable Image Registration for Head and Neck Cancer", Imaging and Computer Assistance in Radiation Therapy Workshop of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention (Munich, Germany, 2015).

N. Burgos, M. J. Cardoso, F. Guerreiro, **C. Veiga**, M. Modat, S. Ourselin, J. McClelland, A.-C. Knopf, S. Punwani, D. Atkinson, S. R. Arridge, B. F. Hutton, and S. Ourselin, "Robust CT Synthesis for Radiotherapy Planning: Application to the Head & Neck Region", Proceedings of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention (Munich, Germany, 2015).

C. Veiga, J. McClelland, S. Moinuddin, K. Ricketts, M. Modat, S. Ourselin, D. D' Souza, and G. Royle, "Towards adaptive radiotherapy for head and neck patients: validation of an in-house deformable registration algorithm," J. Phys.: Conf. Ser. 489 012083 (2014).

C. Veiga, J. McClelland, S. Moinuddin, K. Ricketts, D. D'Souza, and G. Royle, "Deformable registrations for head and neck cancer adaptive radiotherapy", Image Guidance and Multimodal Dose Planning in Radiation Therapy Workshop of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (Nice, France, 2012).

Conference abstracts

A. J. Cole, J. R. McClelland, **C. Veiga**, U. Johnson, D. D'Souza, and M. Bidmead, "Toward adaptive radiotherapy for lung patients: Feasibility study on deforming planning

CT to CBCT to assess the impact of anatomical changes on dosimetry," Proceedings of the 18th International Conference on the Use of Computers in Radiotherapy (London, United Kingdom, 2016).

C. Veiga, G. Janssens, C.-L. Teng, T. Baudier, L. Hotoiu, Lingshu Yin, J. R. McClelland, G. Royle, C.B. Simone II, and B.-K. K. Teo, "Quantitative assessment of proton range deviations using lung CBCT," Proceedings of the 55th Annual Conference Particle Therapy Co-Operative Group (Prague, Czech Republic) (2016).

C. Veiga, T. Long, B. Siow, M. Loizidou, G. Royle, and K. Ricketts, "MO-F-CAMPUS-I-04: Magnetic resonance imaging of an *in vitro* 3D tumor model," Med. Phys. 42(6):3579 (2015).

C. Veiga, J. Alshaikhi, M. Modat, S. Ourselin, G. Royle, R. Amos, and J. R. McClelland, "CBCT and deformable registration based dose calculations for adaptive proton radiotherapy," 4D Treatment Planning Workshop (London, United Kingdom, 2014).

A. M. Lourenço, **C. Veiga**, G. Royle, and J. McClelland, "Dose remapping and summation for head and neck adaptive radiotherapy applications", NPL PPRIG Proton Therapy Physics Workshop (London, United Kingdom, 2014).

C. Veiga, J. McClelland, S. Moinuddin, K. Ricketts, M. Modat, S. Ourselin, D. D' Souza, and G. Royle, "Towards adaptive radiotherapy for head and neck patients: validation of an in-house deformable registration algorithm," Proceedings of the 17th International Conference on the Use of Computers in Radiotherapy (Melbourne, Australia, 2013).

C. Veiga, J. McClelland, S. Moinuddin, K. Ricketts, D. D' Souza, and G. Royle, "Calculation of the dose of the day using an in-house validated deformable registration algorithm," Radiother. Oncol. 106(S2), S478 (2013) (Geneva, Switzerland, 2013).

S. Moinuddin, P. Davies, R. Bodey, **C. Veiga**, R. Mendes, D. D'Souza, G. Royle and I. Rosenberg, "Adaptive re-planning for H/N IMRT: How to choose when to do it!" IPEM: Adaptive radiotherapy (Leeds, United Kingdom, 2013).

Prizes

Runner-up: UCL Graduate School Poster Competition 2013/14 (Built Environment, Engineering Sciences, Mathematical & Physical Sciences), "Image-guided and adaptive

radiation therapy for head and neck cancer" (2014).

Grants

PTCOG Travel Fellowship (2016).

IPEM bursary (2014).

IOP Research Student Conference Fund (2013).

UCL Graduate School Student Conference Fund (2013).

PARSUK Xperience Mentor Grant (2013).

Chapter 1

Introduction

> Living is worthwhile if one can contribute in some small way to this endless chain of progress.
>
> Paul Dirac

1.1 Contextualisation of the research project

Radiation therapy stands for the medical use of ionising radiation as part of cancer treatment. Radiotherapy works by damaging the genetic material of cancerous cells [1]. The treatment is devised such that the prescribed dose is delivered to the tumour while minimising the dose to the surrounding healthy tissues, and delivered over multiple and smaller doses over a period time (fractionation) to minimise the negative side effects of the treatments. Intensity modulated radiation therapy (IMRT) [2, 3] and proton therapy [4] are the state-of-art external radiotherapy modalities. IMRT and proton therapy deliver very precise dose maps that minimise the dose to healthy tissues, and therefore the risks of secondary effects. To make the most of such precise delivery it becomes crucial to have accurate knowledge of the patient anatomy, biology, and tumour response during treatment, as unaccounted variations can compromise the outcome of the treatment. A typical radiotherapy treatment starts with the acquisition of a computed tomography (CT) scan, which is used to plan an individualised treatment for the patient. CT is the universal imaging modality in radiotherapy due to its good image quality, volumetric information, and how it correlates with the dose deposited during the actual treatment. Thus, a radiotherapy treatment is planned on a "snapshot" of the patient, but is actually delivered daily over several weeks, based on the (not always correct) premise that the anatomy is unchanged since the planning stage. During treatment delivery the patient positioning is verified with image guidance techniques. Image-guided radiation therapy (IGRT) is a useful tool that can detect and correct random and systematic change errors

that occur during treatment delivery [5, 6]. Several imaging techniques can be used, such as in-room CT, cone-beam CT (CBCT), magnetic resonance imaging (MRI), ultrasound or planar X-rays. Each technique is associated with costs in terms of machine time and patient imaging dose [7]. By combining different and complementary medical imaging modalities it becomes possible to closely monitor the patient's physical and biological responses throughout the treatment course. This information can then be used to rapidly modify the treatment to take in consideration any changes that could impact the final outcome. This framework is known as adaptive radiation therapy (ART) [8].

1.2 Research question and aims

Even though different imaging modalities provide additional and complementary information of the patient, their further introduction in the radiotherapy pathway is still limited by several reasons. Two of the major challenges in modern radiotherapy are how to combine the information that different imaging modalities at different time points provide in a comprehensive way, and how to use all this information in the best possible way to improve patient outcome.

The answer these questions, this project focuses on two main aims:

1. Development of in-house tools to facilitate the integration of different imaging modalities into ART workflows for clinical investigation. This aim is broken down into the following technical objectives:
    - Investigate and optimise the use of deformable image registration (DIR) for the alignment of CT, CBCT and MRI images, in the context of head and neck (HN) and/or lung malignancies.
    - Proposing and applying methodologies to validate the use of DIR for the clinical applications of contour propagation, dose recalculation and summation.
    - Implementation of the ART workflows as prototype in a research platform tool.
2. Development an *in vitro* tumour model (tumoroid) that can be used for multimodal and sequential imaging studies, and therefore act as a test subject that provides pre-clinical evidence of the benefits of incorporating additional imaging information in radiotherapy. This objective is broken down into the following technical objectives:
    - Design specifications and procedure development to achieve MRI-friendly samples.
    - Design of the experimental setup of pre-clinical MRI sessions of the tumoroids.
    - Preliminary characterisation of the relaxometric properties of the tumoroids.

1.3 My contribution to this work

This research project is in the field of Medical Physics, particularly IGRT and ART, and is part of a collaboration effort across various and interdisciplinary research groups. The following institutions and departments were involved in this collaboration:

- Proton and Advanced RadioTherapy Group, Department of Medical Physics & Biomedical Engineering, London, United Kingdom (UK)
- Centre of Medical Image Computing (CMIC), Department of Medical Physics & Biomedical Engineering and Department of Computer Science, University College London (UCL), London, UK
- Radiotherapy Physics Department, University College London Hospital (UCLH), London, UK
- Radiotherapy Department, UCLH, London, UK
- Centre for Advanced Biomedical Imaging (CABI), UCL, London, UK
- Division of Surgery and Interventional Science, UCL, London, UK
- Ion Beam Applications, SA (IBA), Louvain-la-Neuve, Belgium
- Department of Radiation Oncology, Hospital of the University Pennsylvania (HUP), United States of America (USA)

The different areas of research are listed below, along with my particular contribution to each one:

Clinical needs in radiotherapy: I was part-time based at the Radiotherapy Physics and Radiotherapy departments at UCLH from the beginning of this project to understand the clinical needs of HN and lung patients. This included observing the clinical IMRT treatment pathway, learning about how to use two treatment planning systems (Eclipse and RayStation) for ART applications, and practical sessions on proton treatment planning.

During the time I completed my research project in UCLH multimodal imaging was not used in ART applications, so to fill gaps in knowledge due to lack of in-house expertise, I attended the "Multimodal imaging towards individualized RT treatments" SUMMER consortium summer school in Delft, The Netherlands (July 2014).

Proton beam therapy centres are currently being developed at UCLH and The Christie (Manchester), and will start treating patients from 2018. Therefore, this is still a growing area in the UK and the expertise is still very limited. To further specialise in this area, I attended the NPL PPRIG Proton Workshop in London, UK (March 2014). This workshop was very valuable to learn about the different areas of research in proton therapy, and

was an unique opportunity to understand the need of image guidance in proton therapy directly from experts in the field.

I was also responsible for organising a visit to the Centre of Protonthérapie d'Orsay (Orsay, France) to establish research links with this institution (December 2014).

Finally, I have spent 6 months at Roberts Proton Center at the HUP (Philadelphia, USA) in collaboration with IBA (April to September 2015). I had a privileged role in the development and evaluation of the world's first CBCT system for adaptive lung proton therapy. During my stay I also participated in several clinical activities, including proton therapy patient specific and machine quality assurance.

Computational medical imaging tools: NiftK was the main research tool used in this project, and it was developed by computer scientists at CMIC; NiftyReg is the open-source DIR tool available as part of the NifTK project. My work is in the interface between theoretical/technical developments and clinical usage, and my contributions include applying technologies developed by computer scientists to clinical applications, modifying outputs to a language that can be interpreted by the treatment planning system (TPS) available clinically, implementing an accessible framework that will spark the clinicians' interest, providing validation protocols that answer their concerns, and highlighting the benefits of translating this new technology to the clinic. As part of this process, I was also closely involved in testing cutting-edge improvements of the software tools, identifying and reporting malfunctioning of the code, and assessing its performance and relevance for different applications.

In the context of my placement at the University of Pennsylvania, I was invited to IBA headquarters (Louvain-la-Neuve, Belgium) where I stayed for a week (March 2015) and was introduced to the iMagX project. I was trained to be competent in the in-house DIR tools developed at IBA (REGGUI), which I independently used and further developed during the whole placement.

Tissue Engineering: I was part-time based at the UCL Division of Surgery and Interventional Sciences from September to November (2013) to be trained on the basics of cell culture techniques and tissue engineering. I learnt the protocol for the production of the artificial tumour model, and identified the limitations of the model for multimodal and sequential imaging. I worked very closely with Tong Long (Division of Surgery and Interventional Sciences, UCL) to design a tridimensional (3D) phantom more adequate for such applications. The samples used in the experimental work of this thesis were manufactured by my colleague, but the sample optimisation and imaging experiments were designed by me.

Pre-clinical imaging: I was responsible for setting a collaboration link with CABI to access their pre-clinical MRI scanner for experimental sessions. Dr. Bernard Siow was the responsible for the scanner and for the optimisation of the acquisitions during the experiments. I was responsible for designing and preparing the experimental setup, was trained to independently operate the MRI scanner, and analysed the resulting imaging data.

1.4 Novelty of this work

Research efforts are being focused in developing reliable workflows to incorporate additional imaging as part of the clinical radiotherapy pathway. Several aspects of the work presented in this thesis are novel:

- While the concept of using CBCT and DIR for ART for HN patients itself was first investigated by Yang et al. [9] and Peroni et al. [10], this was the first time the uncertainties associated with dose calculations and dose warping were reported. Particularly, in the context of proton therapy this was, simultaneously with the work of Landry et al. [11, 12], one of the first studies assessing the clinical implications of CBCT and DIR based dose calculations.

- This thesis reports the first clinical use of on-board CBCT for adaptive proton therapy for lung cancer. This included the proposition of a novel adaptive therapy workflow, based on a fast decision online followed by a more careful offline review. This workflow was benchmarked both in terms of clinical indicators generated, and on the uncertainties associated with the approximations used.

- On a more technical aspect, several points of this work were novel. Different gold-standards for the validation of CT-to-CBCT registration were proposed. Additionally, NiftyReg had not been previously validated in the HN region for different image modalities (CT, CBCT, and MRI), or validated specifically for ART applications. A new method to deal with missing image information for CBCT dose calculations was proposed, which was adequate for the HN region. Finally, a correction method was employed to deal with the limitations of deformable registration regarding non-deformable changes the thoracic region.

- The idea of developing a tumour model tailored for MRI is novel, and so is the tumour model engineered toward this application. This was the first attempt to acquire images of the artificial cancer masses on pre-clinical MRI system and to quantify its relaxometric properties.

The work presented in this thesis resulted in the following peer-reviewed journal papers:

- *(In preparation)* C. Veiga, G. Janssens, T. Baudier, L. Hotoiu, S. Brousmiche, J. R.

McClelland, C.-L. Teng, L. Yin, G. Royle, and B.-K. K. Teo, "The accuracy of CBCT and deformable registration for adaptive lung proton therapy" (2016).

- C. Veiga, G. Janssens, C.-L. Teng, T. Baudier, L. Hotoiu, J. R. McClelland, G. Royle, L. Lin, L. Yin, J. Metz, T. D. Solberg, Z. Tochner, C. B. Simone II, J. McDonough, and B.-K. K. Teo, "First clinical investigation of CBCT and deformable registration for adaptive proton therapy of lung cancer," Int. J. Radiat. Oncol. Biol. Phys. 95(1) 549-559 (2016).
- C. Veiga, J. Alshaikhi, R. Amos, A. M. Lourenço, M. Modat, S. Ourselin, G. Royle, and J. R. McClelland, "CBCT and deformable registration based "dose of the day" calculations for adaptive proton therapy," Int. J. Particle Ther. 2(2) 404-414 (2015).
- C. Veiga, A. Lourenço, S. Moinuddin, M. van Herk, M. Modat, S. Ourselin, D. D'Souza, G. Royle, and J. R. McClelland, "Toward adaptive radiotherapy for head and neck patients: uncertainties in dose warping due to the choice of deformable registration algorithm," Med. Phys. 42(2) 760-769 (2015).
- C. Veiga, J. McClelland, S. Moinuddin, A. Lourenço, K. Ricketts, J. Annkah, M. Modat, S. Ourselin, D. D'Souza, and G. Royle, "Toward adaptive radiotherapy for head and neck patients: feasibility study on using CT-to-CBCT deformable registration for "dose of the day" calculations," Med. Phys. 41 031703 (2014).

1.5 Impact of this work

The work conducted was truly collaborative and multidisciplinary; thus it had impact beyond the content explicitly exhibited in this thesis:

Clinical tools for DIR: The tools developed throughout this project were implemented in a friendly way in clinical research settings, and are currently being used at the Department of Radiotherapy Physics (UCLH) and Radiation Oncology (HUP) to monitor patients that may benefit from treatment adaptation.

Adaptive lung therapy for photon therapy: Following the studies performed on the context of HN malignancies for photon/proton therapy at UCLH and for lung at HUP, I collaborated with Alison Cole (Department of Radiotherapy Physics, UCLH) to extend the work performed on HN to lung malignancies in the context of photon therapy at UCLH. This work resulted in the following output:

- A. J. Cole, J. R. McClelland, C. Veiga, U. Johnson, D. D'Souza, and M. Bidmead, "Toward adaptive radiotherapy for lung patients: Feasibility study on deforming planning CT to CBCT to assess the impact of anatomical changes on dosimetry,"

Proceedings of the 18th International Conference on the Use of Computers in Radiotherapy (London, UK, 2016).

Range verification for eye proton therapy based on proton-induced x-ray emissions from implanted metal markers: Due to my computational skills and expertise of proton therapy, I was part of a team consisting of members from UCL and Centro de Adroterapia e Applicazioni Nucleari Avanzate (CATANA) proton source at the Istituto Nazionale Fisica Nuclear - Laboratori Nazionali del Sud (INFN-LNS) (Catania, Italy), which operated to collect experimental data in April 2013 [13].

Validation of clinical acceptability of an atlas-based segmentation algorithm for the delineation of organs at risk in head and neck cancer: I collaborated with Albert Duanc (CMIC, UCL) to facilitate the transfer of data between NifTK and the clinical systems for validation of automatic segmentation in the HN, and provided expertise into the validation of the application for clinical use. This work resulted in the following output:

- A. K. Hoang Duc, G. Eminowicz, J. McClelland, M. Modat, M. J. Cardoso, A. F. Mendelson, C. Veiga, T. Kadir, D. D'Souza, and S. Ourselin, "Validation of clinical acceptability of an atlas-based segmentation algorithm for the delineation of organs at risk in head and neck cancer," Med. Phys. 42(9) 5027-5034 (2015).

Synthesising CT from MRI data: I collaborated closely with Dr. Albert Duanc (CMIC, UCL) on a synthetic CT project based on atlas registration, and ran preliminary analysis on the clinical performance of the method [14]. The preliminary results obtained were promising, and therefore a collaboration group was formed with the Institute of Cancer Research, that within the MRI-linear accelerator (LINAC) project was developing treatment planning on MRI. I provided expertise in the optimisation of DIR for registrations of images of the HN region, and how to clinically validate the dose calculations. This work resulted in the following output:

- N. Burgos, M. J. Cardoso, F. Guerreiro, C. Veiga, M. Modat, S. Ourselin, J. McClelland, A.-C. Knopf, S. Punwani, D. Atkinson, S. R. Arridge, B. F. Hutton, and S. Ourselin, "Robust CT Synthesis for Radiotherapy Planning: Application to the Head & Neck Region", Proceedings of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention (Munich, Germany, 2015).

Six degrees-of-freedom couch quality assurance in clinical settings: I provided expertise in image registration and research/commercial system integration to generate synthetic rotated phantoms for quality assurance of the new six degrees-of-freedom couch at the Radiotherapy department, UCLH. This is now being used clinically.

Development of oxide contrast agents: The tools developed to quantify the relaxometric properties of the tumour models from the data extracted from the pre-clinical magnetic resonance (MR) system at CABI were modified for other application. This project was conducted by Nicholas Hobson and Xian Weng Jiang (School of Pharmacy, UCL) and consisted of developing oxide contrast-agents for cancer detection and drug delivery. This work resulted in the following outputs:

- *(In preparation)* N. Hobson, X. Weng-Jiang, C. Veiga, B. Siow, M. Ashford, N. T. K. Thanh, A. Schätzlein, and I. Uchegbu, "Design and synthesis of self-assembling polymeric iron oxide nanoparticle theranostics for applications in cancer diagnostics and cancer therapy," ACS Nano (2016).
- *(In preparation)* X. Weng-Jiang, N. Hobson, C. Veiga, B. Siow, N. T. K. Thanh, A. Schätzlein, and I. Uchegbu, "Aqueous in-flow synthesis of superparamagnetic iron oxide nanoparticles for dual T_1/T_2-weighted magnetic resonance imaging," ACS Nano (2016).

1.6 Structure of this thesis

The current chapter consisted of a brief introduction to the research context where this thesis is inserted, detailing the research questions, novelty and personal contribution to the work. The structure of the remaining of the thesis resulted from grouping its contents in two major lines of research, described and justified in the following paragraphs.

The first line of research of this thesis, consisting of chapters 2 to 7, presents very focused and coherent studies into the same common global topic of the use and validation of DIR in ART applications. This research line was motivated by clinical needs from the Radiotherapy Department at UCLH. Chapter 2 introduces the theoretical concepts of cancer radiobiology, image registration, and the clinical problems tackled in this context. It also includes all the preliminary work conducted regarding the optimisation of DIR, and the identification of strategies for its validation for clinical applications. Building on this introduction, the following chapters evaluate the use of image registration for different and sequential applications. Chapter 3 is focused on the use of DIR for contour propagation and "dose of the day" calculations for HN malignancies and in the context of IMRT treatments. Then, this work was extended to the study of dose warping and summation in chapter 4, and to proton therapy in chapter 5. The work of these three chapters builds up the expertise that culminates in chapter 6, where a clinical adaptive therapy workflow based on CBCT and DIR was implemented, thoroughly validated and clinically investigated in the context of lung proton therapy. Finally, in chapter 7 preliminary work on co-registration of multimodal (CT and MR) imaging was reported. The focus on MRI instead of CBCT on this chapter creates a bridge with the work presented in the following chapter.

The second line of research consists of the final chapter (chapter 8), in which the development of an *in vitro* artificial cancer mass (ACM) for multimodal and multitemporal imaging experiments was investigated. The methods and materials used in this chapter differ substantially from those used in the previous chapters (i.e., computational methods and patient data versus experimental methods and *in vitro* data). In the previous chapters, two of the major technical difficulties found while conducting the studies presented was the limitations of the readily available clinical data (both in acquisition protocols and size of the cohorts), and the difficulties in defining ideal gold-standards for DIR validation based on patient models. Those two points are the reason why this study is presented at the end of this thesis, and in conjunction with excellent collaboration links with the UCL Division of Surgery, constitute the motivation behind the use of controllable *in vitro* data.

Chapter 2

The role of deformable image registration in adaptive radiotherapy

> I learned very early the difference between knowing the name of something and knowing something.
>
> Richard Feynman

This chapter introduces the role of image registration in ART and NifTK, the main software tool used in the project. The word described here allowed to identify strategies for the validation of DIR for radiotherapy applications, as well as developing of the tools needed for its in-house clinical translation.

The work in this chapter resulted in the following outputs:

- C. Veiga, J. McClelland, S. Moinuddin, K. Ricketts, M. Modat, S. Ourselin, D. D' Souza, and G. Royle, "Towards adaptive radiotherapy for head and neck patients: validation of an in-house deformable registration algorithm," J. Phys.: Conf. Ser. 489 012083 (2014).

- S. Moinuddin, P. Davies, R. Bodey, C. Veiga, R. Mendes, D. D'Souza, G. Royle and I. Rosenberg, "Adaptive re-planning for H/N IMRT: How to choose when to do it!" IPEM: Adaptive radiotherapy (Leeds, United Kingdom, 2013).

- C. Veiga, J. McClelland, S. Moinuddin, K. Ricketts, D. D'Souza, and G. Royle, "Deformable registrations for head and neck cancer adaptive radiotherapy", Image Guidance and Multimodal Dose Planning in Radiation Therapy Workshop of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (Nice, France, 2012).

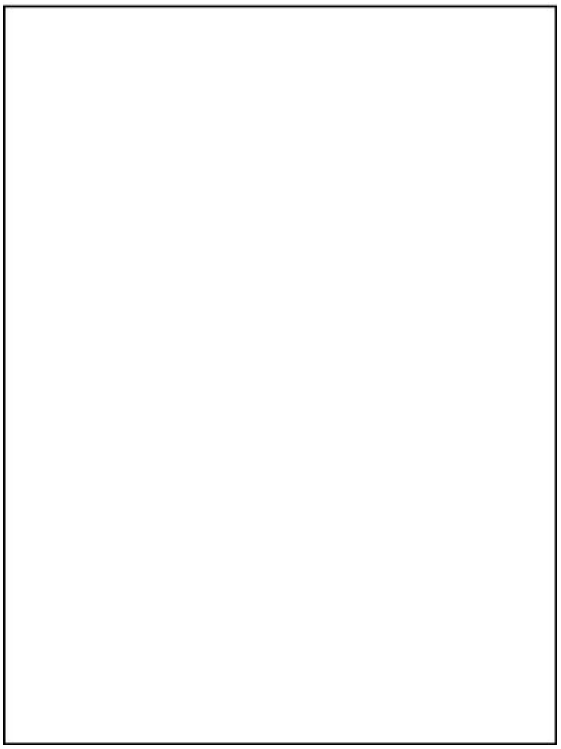

Figure 2.1: *Direct and indirect actions of radiation [1].*

2.1 An introduction to cancer radiobiology

Cancer begins when a cell breaks free from the normal restraints on cell division and begins to follow its own agenda for proliferation [15]. In order to fully understand cancer treatment using radiotherapy, an understanding of the biology of the tumour microenvironment and biological effects of radiation is necessary.

The biological effects of radiation result principally from damage to the deoxyribonucleic acid (DNA), which is the critical target in radiotherapy. The radiation is known to interact in two distinct pathways: (1) direct action, i.e., the radiation interacts directly with the DNA molecule, and (2) indirect action, i.e., the radiation interacts with the water inside the cell producing free radicals that interact with the DNA (Figure 2.1). About two-thirds of the biological damage caused by x-rays results from indirect action. The timeline of physical/chemical and biological effects are of very different orders of magnitude. The physics of the absorption process is 10^{-15}s; the chemistry takes 10^{-5}s for the reactions between DNA and free radicals; the biology takes hours, days or months for cell killing [1]. In radiobiology, cell death is defined as the process that leads to permanent loss of reproductive capacity which includes several mechanisms. The most common process in radiotherapy is mitotic death, but other mechanisms such as apoptosis, autophagy, necrosis, and senescence are also possible responses [16]. The prevalence of each mechanism differs between different types of normal and tumour cells.

The tumour microenvironment consists of the cancer cells, normal cells, structural

support given by the extracellular matrix (ECM) and secreted soluble factors that regulate the growth and signalling between cells. Hence, the complexity of the *in vivo* system causes a non linear relationship between physical dose and biological effect. The biological consequences of DNA damage are complex and influenced by pathways within the DNA damage response system, which determines the likelihood of the cells dying after irradiation and the type of cell death that occurs. Depending on the severity of the damage caused by a single irradiation, this damage may be irreversible and irreparable, leading to cell death (i.e., lethal damage). However, if the damage is not lethal the cells have mechanisms of DNA repair and may be able to recover for sublethal damage. Tumour cells are known to have lost the ability of repair damage, and thus are in general more sensitive to irradiation than normal tissue. Moreover, biological effects that are not related with direct dose delivery also occur, as irradiated cells signal nearby unirradiated cells that also exhibit response to radiation (bystander effect) [17].

Several factors are known to influence the radiosensitivity of human cells (i.e., the biological outcome will differ when the same physical dose is delivered):

- Cell cycle: cells are more sensitive to radiation depending on the phase of their cell cycle (Figure 2.2) [18]. Between cell cycle phases checkpoints exist, such that damaged normal cells stop progressing through the cycle to attempt to repair damage. Abnormalities in the genetic material of cancer cells interfere with the repair mechanisms, and the cells progress to mitosis anyway leading to mitotic catastrophe. In general, cells are more sensitive to radiation during the mitotic phase as repair is not possible at this point.

- Oxygenation: aerobic cells are generally more radiosensitive than hypoxic cells;

- Proliferation: the higher the rate of proliferation, the greater the radiosensitivity;

- Differentiation: undifferentiated cells are more radiosensitive than differentiated cells.

Therefore, and considering the biological mechanisms of dose response, two parameters are of utmost importance when devising a radiotherapy treatment: dose rate and fractionation. Radiation-induced cell death is directly proportional to dose rate. However, both normal and tumour cells show this increased radiosensitivity, hence high dose rates are rarely used to improve radiotherapy outcome. Dose is then delivered in fractions for the following reasons: (1) it allows for re-oxygenation of previously hypoxic tumour areas (Figure 2.3), (2) permits the redistribution of cells in the cell cycle, increasing the proportion of cancer cells in more radiosensitive phases of the cell cycle on the next radiotherapy fraction; and (3) normal cells exhibited higher rates of repair than tumour cells, and hence are given time to recover from radiation damage and repopulate. The schemes of fractionation used in the clinic are based on empirical data and convenience [1, 19].

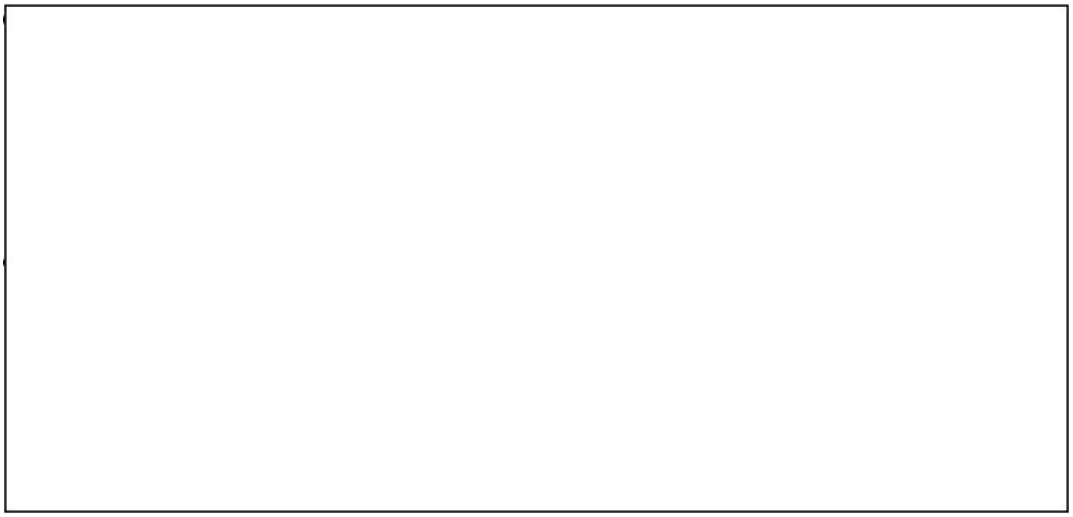

Figure 2.2: *(a) Cell cycle phases: gap 0 (G_0), gap I (G_1), synthesis (S), gap II (G_2) and mitosis (M). (b) Variation of radiosensitivity with the phase of the cells in the cell cycle [18].*

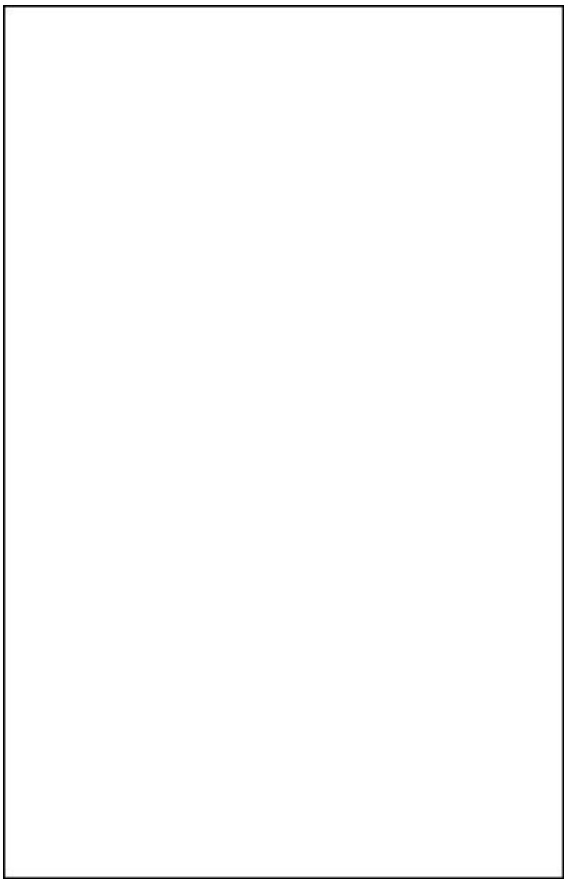

Figure 2.3: *Tumours contain a mixture of aerated and hypoxic cells. Irradiation kills a greater fraction of aerated than hypoxic cells, leaving mostly hypoxic cells surviving. Given time reoxygenation occurs, and the distribution of aerated/hypoxic cells returns to pre-irradiation state. This allows to successfully target previously radioresistant hypoxic cells. Adapted from [1].*

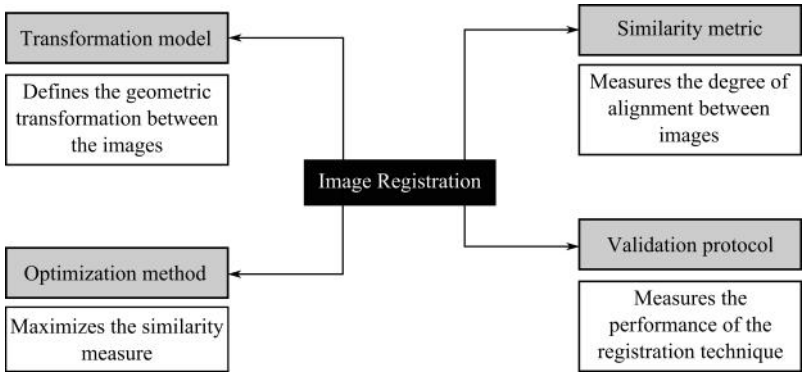

Figure 2.4: *Key items of any image registration algorithm: transformation model, similarity metric, optimisation method and validation protocol [20].*

2.2 An introduction to image registration

The ability to fully utilise the information extracted from images acquired with different modalities and/or at different time-points relies on the accuracy to align the multiple sources of information. Image registration is the process of aligning different sets of data into a single coordinate system. It can be used to find the corresponding anatomical, biological or functional locations between two or more images. Depending on the specific application, it can be divided based on if the subject is the same or not for the different images (intra-subject or inter-subject registration), or by taking in consideration if the images being registered are of the same modality or not (monomodal or multimodal registration).

The key items of an image registration algorithm are summarised in Figure 2.4, and will be discussed in further detail in the following sections.

2.2.1 Transformation Model

Image registration results in a mathematical transformation **T** that maps every point in a source (or floating) image to the corresponding point in a target (or reference) image.

$$\mathbf{T} : (x, y, z) \rightarrow (x', y', z') \tag{2.1}$$

There are several transformation models, ranging from quite simple, like rigid and affine transformations, to more complex, like deformable transformations.

A rigid transformation in three dimensions involves the rotation and translation in the three different Cartesian axes. An affine transformation combines rigid alignment with scaling and shearing. Rigid and affine transformations are usually applied in the registration of anatomical structures like the brain and bones. However, when significant deformation is expected, like in soft tissue, such simple transformations do not properly

characterise the deformation of the tissue. In situations like this non-rigid (or deformable) transformations are used. There are a plethora of DIR algorithms available in the literature, which can be divided into parametric methods (B-splines [21], linear elastic finite element method [22], etc.) and nonparametric methods (viscous fluid [23], Demons [24], etc.).

Free-form deformations (FFDs) are a popular type of DIR algorithm. The basic idea is to deform an object by deforming the space around it, that is, by manipulating a 3D parallelepiped lattice containing the object. This manipulated lattice determines a deformation function that specifies a new position for each point in the object [25]. The original FFD scheme was based on trivariate Bernstein polynomials [26], but tri-variate B-splines tensor products are used nowadays [25, 27]. The use of FFD based on B-splines was first proposed by Rueckert et al. for the registration of contrast-enhanced breast MRI [21]. Spline-based transformations are based on the assumption that a set of corresponding control points (CPs) can be identified in the source and target images. At the CP position the spline-based transformations interpolate or approximate the displacements, which are necessary to map the location of the CP in the target image to its corresponding counterpart in the source image. Between CPs, they provide a smoothly varying displacement field. The local control properties of B-splines make them computationally efficient even for a large number of CPs, and the continuity of the transformation is guaranteed when any CPs are moved [28], as only the local neighbourhood of that CP is affected [21].

To define a B-spline based FFD, the domain of the image volume is defined as $\Omega = \{(x, y, z)|0 \leq x < X, 0 \leq y < Y, 0 \leq z < Z\}$, and Φ is a $n_x \times n_y \times n_z$ mesh of control points $\phi_{i,j,k}$ with spacing δ. Then, the FFD can be written as a 3D tensor of the unidimensional (1D) cubic B-splines.

$$\mathbf{T}_{\text{local}} = \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3} B_l(u)B_m(v)B_n(w)\phi_{i+l,j+m,k+n} \tag{2.2}$$

where $i = \lfloor x/\delta_x \rfloor - 1$, $j = \lfloor y/\delta_y \rfloor - 1$, $k = \lfloor z/\delta_z \rfloor - 1$, $u = x/\delta_x - \lfloor x/\delta_x \rfloor$, $v = y/\delta_y - \lfloor y/\delta_y \rfloor$, $w = z/\delta_z - \lfloor z/\delta_z \rfloor$ and B_l is the lth basis function of the B-spline [25, 27].

$$B_0(u) = (1-u)^3/6 \tag{2.3a}$$

$$B_1(u) = (3u^3 - 6u^2 + 4)/6 \tag{2.3b}$$

$$B_2(u) = (-3u^3 + 3u^2 + 3u + 1)/6 \tag{2.3c}$$

$$B_3(u) = u^3/6 \tag{2.3d}$$

2.2.2 Similarity metric

The registration looks to find correspondences of voxel intensities in the field-of-view (FoV) of the images, and the algorithm will maximise some measure of similarity. The similarity metric measures globally or locally the degree of alignment between the images

registered (i.e., how well the images are matched to each other). For monomodal registrations, images of similar histogram content are registered by establishing a relationship between pixel intensities, while for multimodal registrations the assessment of pixel similarities is replaced by the likelihood of a pixel position being occupied [29]. Some of the most popular measures of similarity are described below:

- Sum of the squared differences (SSD)

$$\text{SSD} = \frac{1}{N}\sum_{i}^{N}(A(i) - B(i))^2 \tag{2.4}$$

  where N is the number of voxels in the region of overlap.

- Normalised cross correlation (NCC)

$$\text{NCC} = \frac{\sum(A(i) - \bar{A})(B(i) - \bar{B})}{\sqrt{\sum(A(i) - \bar{A})^2 \sum(B(i) - \bar{B})^2}} \tag{2.5}$$

  where $\bar{A}$ and $\bar{B}$ are the average intensities of the two images.

- Normalised mutual information (NMI), which is based on the information content, or entropy, of the images:

$$\text{NMI(A,B)} = \frac{H(A) + H(B)}{H(A, B)} \tag{2.6}$$

  The entropy $H(A)$ of an image A is:

$$H(A) = -\sum_{a \in A} p(a) \log p(a) \tag{2.7}$$

  where $p(a)$ is the probability that a voxel in image A has intensity a. The joint entropy $H(A, B)$ of the overlapping region of images A and B is

$$H(A, B) = -\sum_{a \in A}\sum_{b \in B} p(a, b) \log p(a, b) \tag{2.8}$$

  where $p(a, b)$ is the joint probability that a voxel in the overlapping region of A and B has values a and b.

Both SSD and NCC assume that both image modalities have the same intensity characteristics. If the images are correctly aligned, the different between them should be zero except for the noise produced. Such measures can be calculated globally (i.e., over the whole common FoV) or localised (i.e., over a specified neighbourhood of the pixel). An example of such a measure is the localised normalised cross correlation (LNCC).

NMI is based on the notion of the marginal and joint probability distributions of the two images, and therefore is adequate for multimodal applications. It can be estimated by using histograms whose bins count the frequency of occurrence (or co-occurrence) of intensities. Dividing these frequencies by the total number of voxels yields the estimate of the probability of that intensity.

2.2.3 Optimisation

Optimisation in image registration aims at (1) maximising the similarity of the images and (2) minimising the cost associated with particular transforms. Thus, a cost function is defined as the sum of the measure of similarity and constrains, added to stop the registrations from being ill-posed (i.e., having no unique or stable of solution). The constraints act as a regularisation of the transformation.

In clinical applications, it is commonly accepted that the local deformation of soft tissue should be characterised by a smooth transformation. A B-spline FFD can be constrained to be smooth by introducing a 3D penalty term that regularises the transformation. A popular constrain is the bending energy penalty term (BE) [21]:

$$\mathrm{BE} = \int\int\int_{\Omega}\left[\left(\frac{\partial^2\mathbf{T}}{\partial x^2}\right)^2+\left(\frac{\partial^2\mathbf{T}}{\partial y^2}\right)^2+\left(\frac{\partial^2\mathbf{T}}{\partial z^2}\right)^2+2\left[\left(\frac{\partial^2\mathbf{T}}{\partial xy}\right)^2+\left(\frac{\partial^2\mathbf{T}}{\partial xz}\right)^2+\left(\frac{\partial^2\mathbf{T}}{\partial yz}\right)^2\right]\right]$$

Other penalty terms can be used to constrain the registrations, such as the logarithm of the Jacobian determinant penalty term (JL):

$$\mathrm{JL} = \frac{1}{n}\sum|\log(\det(\nabla\mathbf{T})| \tag{2.9}$$

The Jacobian determinant has an important physical meaning: det(Jac)=1 means that there is no volume change, while det(Jac)<1 is a compression and det(Jac)>1 an expansion. Negative det(Jac) is in general unwanted since it means that the pixel disappears (a process also known as folding). JL penalises the regions where the algorithm tries to do extreme contractions or expansions, and enforces one-to-one mapping in the resulting transformation.

The cost function is maximised using an optimisation algorithm. Popular choices of algorithms are the gradient descend and conjugate gradient methods. Such algorithms require the gradient of the cost function, so faster implementations are possible when this can be calculated analytically.

2.2.4 Symmetry, inverse-consistency and diffeomorphisms

In image registration at least two input images are necessary, the source and the target images, and the result is a transformation ($T_{s\to t}$) that can be used to deform the source image (s) onto the target image (t). Therefore if $T_{s\to t}$ exists, the transformation in the opposite direction ($T_{t\to s}$), which can be used to deform the target image onto the source image, can also be defined. The majority of the research and commercial registration algorithms are unidirectional, which means they only optimise and generate the forward transformation ($T_{s\to t}$) and do not consider the transformation in the opposite direction

($T_{t\to s}$). For applications where $T_{t\to s}$ is also required one can simply use a unidirectional algorithm twice, by switching the roles of the source and the target images, numerically estimate the inverse of $T_{s\to t}$, or use bidirectional algorithms that optimise $T_{s\to t}$ and $T_{t\to s}$ simultaneously.

In clinical applications, DIR can be used to model the spatial anatomical mapping between time points, therefore physically plausible transformations may be desirable for applications where the underlying deformation is important. Two concepts are associated with physically plausible deformations: inverse-consistency and symmetry. Inverse-consistent registrations try to ensure that $T_{t\to s}$ is the mathematical inverse of $T_{s\to t}$ (i.e., $T_{t\to s} = T_{s\to t}^{-1}$). Symmetric registration means that identical transformations are obtained when the roles of source and target images are switched: if the source image becomes the target (t', such that $t' = s$), and the target is now the source (s', such that $s' = t$), a symmetric algorithm will ensure that $T_{t'\to s'}$=$T_{s\to t}$ and $T_{s'\to t'}$=$T_{t\to s}$. The two concepts are usually intertwined in the literature, but are not equal as a symmetric algorithm is not necessarily inverse-consistent (i.e., $T_{s'\to t'}$ is not guaranteed to be $T_{t'\to s'}^{-1}$), and vice-versa. The differences between symmetry and inverse-consistency are more clear when considering unidirectional algorithms, since most bidirectional algorithms that aim to guarantee inverse-consistency are also symmetric. For example, when performing two unidirectional registrations, one in each direction, the resulting transformations are symmetric but not inverse-consistent (switching the source and target results in the same the transformations on opposite directions). If $T_{t\to s}$ is obtained by numerically estimating the inverse of the final registration result ($T_{s\to t}$) the transformations are inverse-consistent, but are not symmetric (estimating the inverse of the opposite transformation does not produce the same result as running the unidirectional registration in that direction).

Another commonly stated requirement for DIR is to have diffeomorphic transformations, i.e., deformations that are invertible, differentiable and whose inverse is also differentiable [30]. Diffeomorphic transformations maintain the topology and guarantee that connected subregions of an image remain connected, neighbourhood relationships between structures are preserved, and surfaces are mapped to surfaces [31]. A diffeomorphic transformation implies invertibility in the sense that T^{-1} is defined. However not all diffeomorphic registration algorithms explicitly generate T^{-1}.

In recent years advanced and complex registration algorithms have been developed to be symmetric, inverse-consistent and diffeomorphic. One approach consists in using the inverse-consistency error (ICE) to create a inverse-consistency penalty term (IC) [31]:

$$\mathrm{IC} = \sum_{\mathbf{x}\forall R} \|\mathbf{T}_{Fw}(\mathbf{T}_{Bw}(x))\|^2 + \sum_{\mathbf{x}\forall F} \|\mathbf{T}_{Bw}(\mathbf{T}_{Fw}(x))\|^2 \tag{2.10}$$

While this encourages inverse-consistency, it can reduce the ability to recover large and complex deformations, and the forward and backward transformations are only

approximate inverses to each other [32]. A better, but less add-hoc, approach is to use a stationary velocity field to parametrise the transformation [33]. In large deformation models the displacement field $\mathbf{u}$ is generated via a time dependent velocity field,

$$\mathbf{u}(x, y, z, 1) = \int_0^1 \mathbf{v}(\mathbf{u}(x, y, z, t))dt \tag{2.11}$$

with $\mathbf{u}(x, y, z, 0) = (x, y, z)$. This can be used to generate the deformation field in either the forward or backward direction, and these are guaranteed to be exact inverses of each other (subject to methodological approximation and numerical precision). The resulting transformation provides a smooth one-to-one (invertible) mapping.

2.2.5 Evaluation and validation of deformable image registration

Evaluation and validation are two steps of major importance to image registration; however, this is not incorporated as part of the software tools and in-house strategies are necessary. Evaluation and validation are necessary to quantify the performance and to show the suitability of the algorithms for the desired medical application. There are many criteria to take in consideration for the evaluation and validation of an image registration technique. These can involve the following and more [20]:

- technical criteria: is the technique fast, robust, accurate and reliable?
- application criteria: is the technique user-friendly and useful for daily clinical practice?
- legal criteria: does the technique have potential to be used in commercial systems?

2.2.6 In-house software: NifTK

NifTK software was developed by CMIC, at the Department of Medical Physics & Bioengineering of UCL (http://cmic.cs.ucl.ac.uk/home/software). The software is an ongoing project, with regular updates and new functionalities being implemented, and contains several tools for image registration and visualisation. It combines a set of different toolkits, including the open-source NiftyReg for rigid and deformable registration (Table 2.1), and a viewer (NiftyView).

The affine registration implemented in NiftyReg uses a Block Matching-based approach [34]. The default DIR is a standard unidirectional graphics processing unit (GPU) implementation of the popular B-spline FFD algorithm using NMI as similarity measure [21]. The major differences from the original work by Rueckert et al. are in the calculation of the gradient and joint histogram, introduction of other penalty terms, and the GPU implementation, which make the code faster and more robust [35]. Recently more sophisticated (bidirectional, symmetric and inverse-consistent) implementations have been incorporated in the software. An inverse-consistent symmetric [36] and a stationary

Table 2.1: *Functionalities implemented in NiftyReg.*

Command	Tools implemented
reg_aladin	Rigid/affine registration algorithm
reg_f3d	Deformable image registration algorithm
reg_jacobian	Computation of the Jacobian matrix, Jacobian determinant and logarithm of the jacobian determinant of a deformation field
reg_resample	Resampling of images after applying input transformations
reg_transform	Conversion between control points and deformation/displacement fields
	Composition of transformations
	Inversion of affine transformations and deformation fields
	Update the header of an image to incorporate affine transformation

velocity field transformation model implementations [37] became more recently freely available. NiftyReg also features a numerical estimation of the inverse of a deformation vector field (DVF), which uses an iterative method to estimate each vector of the inverse DVF independently using the simplex algorithm. It is similar to other published implementations [38, 39] but was independently developed and implemented by Dr Marcel van Herk (Netherlands Cancer Institute) for this particular project. This algorithm was also implemented in the GPU.

The different algorithms implemented in NiftyReg have similar implementations with different underlying desirable properties. While in the standard and symmetric implementations the transformation at the CPs are directly optimised, in the stationary velocity fields the control points are used to parametrise a stationary velocity field, from which the final transformation is computed through exponentiation.

NiftyReg allows its users the flexibility of defining the parameters of the DIR algorithm. This allows to fully customise the registrations to the particularities of the datasets and application. Table 2.2 describes the most relevant parameters the user is free to tune when performing DIR. SSD, NMI and LNCC are available as similarity measures; BE, JL and IC are some of the available penalty terms.

2.2.6.1 Data transfer between NifTK and clinical systems

Clinical imaging datasets are formatted in the form of Digital imaging in communications in medicine (DICOM). DICOM is a standard protocol to handle, store, print and transmit information in medical imaging and it includes a file format definition (.dcm extension) [40]. A DICOM file saves not only the pixel matrix that forms the actual image, but also information that uniquely defines its origin (i.e., scanner, institution, operator, patient information).

Table 2.2: *Deformable image registration parameters that the user can specify in NiftyReg.*

Flag	Definition
	Inputs and outputs
-ref	Filename of the reference image
-flo	Filename of the floating image
-cpp	Filename of the output control point grid
-res	Filename of the resampled image (registration result)
-rmask	
-fmask	Filenames of the mask image in the reference/floating space
-aff	Filename which contains an input affine transformation
	Input image options
–rLwTh	
–fLwTh	Lower threshold to apply to the reference/floating image
–rUpTh	
–fUpTh	Upper threshold to apply to the reference/floating image
	B-Spline options
-sx	
-sy	
-sz	Grid spacing at highest resolution level along the x/y/z axes
-sym	Use the symmetric implementation of the B-Spline algorithm*
-vel	Use the diffeomorphic implementation of the B-Spline algorithm*
	Regularisation options
-be	Weight of the bending energy penalty term
-jl	Weight of the logarithm of the Jacobian determinant penalty term
-ic	Weight of the inverse-consistency penalty term*
	Similarity measures
–rbn	Number of bin to use for the reference/floating image histogram
–fbn	(normalised mutual information is the default similarity measure).
–lncc	Define localised normalised cross correlation as similarity measure,
	and the standard deviation of the Gaussian kernel*
–ssd	Define sum of the squared differences as similarity measure
Optimisation options	
-maxit	Maximum number of iterations per level
-ln	Number of resolution levels
-lp	Number of the first resolution levels that will be performed

* These parameters were only implemented until more recent versions of the software became available and therefore were not fully investigated until chapters 4 and 5.

NifTK uses as default Nifti format to communicate data, a simplified imaging format developed for research purposes. It is therefore necessary to convert DICOM to Nifti to use the registration package, and NiftyView had the tools necessary for this. The opposite conversion, Nifti to DICOM is more complex as DICOM stores more of information than Nifti, and was not available in the NifTK package. As an auxiliary and necessary step for this project, I wrote MATLAB code to convert images, structures and doses from Nifti to DICOM format that were compatible with the UCLH clinical TPS. This technical work proved to be very useful to other projects I contributed to outside the scope of this thesis.

2.2.7 Other registration algorithms

Even though most of the work conducted in this thesis used NiftyReg as DIR tool, in chapter 6 other DIR softwares were also investigated: REGGUI and the commercial version of DIR in RayStation (RaySearch, Stockholm, SE) TPS. The REGGUI DIR package uses a diffeomorphic Morphons algorithm, a nonparametric method using a phase-based approach. The principle of the method is to match transitions (between dark and bright zones) by looking locally at the spatial oscillations in intensities. This method uses Gaussian smoothing as regularisation of the displacement field and additive accumulation during the iterative process [41]. The DIR algorithm in RayStation 4.5 is hybrid free-form registration that uses a multi-resolution approach and an intensity based similarity measure.

2.3 The role of image registration in image guidance and adaptive radiotherapy

One of the main challenges in clinical radiotherapy is to position the patient in every fraction of the treatment exactly as he was imaged for planning, and make sure he stays in that position during the beam-on time. The steeper the dose gradients, the more important it becomes to precisely position the patient and its internal anatomy [42]. The clinical consequences of inaccuracy include both potential underdosage of the target volumes (resulting in increased risk of tumour recurrence) and potential overdosage of normal tissues (resulting in increased risk of complications) [43]. Setup errors, inter- and intra-fraction organ motion/deformation should be characterised, controlled, and taken into account [44]. IGRT is a useful tool that can detect and correct random and systematic errors that occur during treatment delivery.

The current paradigm of a radiotherapy treatment starts with the acquisition of a CT scan, which is used to plan individualised treatment for the patient. The treatment is then delivered in many fractions over several weeks, based on the premise that the anatomy is unchanged since the planning stage. However, it is well known that patient's anatomy

can vary within a fraction, with swallowing and respiratory motion [45], and from fraction to fraction, with changes in bladder/bowel filling and tumour shrinkage [46]. The concept of ART was first suggested by Yan et al., and can be described as a closed-loop feedback process that suggests a change of paradigm in radiotherapy [8]: the stationary anatomy is replaced by a variable anatomy, by utilising daily imaging in the radiotherapy process [47].

ART is a very broad subject and full clinical implementation requires further developments in computational power, image guidance, dose verification and plan adaptation [47, 48]. Repetitive and daily imaging plays a vital role in ART, and the information retrievable depends on the systems available in the clinic. Of all the systems available, CBCT is an increasingly popular in-room imaging method that provides valuable 3D information of the patient in treatment position. The image quality of CBCT is consistently inferior to CT in soft tissue contrast, Hounsfield unit (HU) consistency and artefacts [47, 49, 50]. For example, the CBCT values vary with the size of the imaged volume, and in result the image intensities in the upper thorax area are in general lower than in the neck area [11]. The lower quality of CBCT imaging limits its direct utilisation to assess the current plan and modify it if necessary.

It is widely accepted that the future of ART depends on the use of DIR algorithms [51–53]. DIR provides a solution for the major challenges in ART: the planning computed tomography (pCT) can be deformed to match the daily anatomy (from in-room imaging, such as CBCT) to calculate the dose delivered per fraction [9], the deformations can be applied to propagate contours [10, 54], and the fraction by fraction dose maps can be warped back to a common reference frame for summation [55, 56].

2.3.1 The clinical problem: head and neck

The definition of HN malignancies covers a heterogeneous group of cancers, which includes paranasal and sinonasal cancer, and cancer of the salivary gland, lip, oral cavity, pharynx and larynx [57]. The annual incidence of the cancer of the oral cavity, pharynx and larynx is approximately 147500 in Europe. This represents 4.6% of all cancer cases, accounting for 63400 deaths [58]. It is considered a complicated cohort to treat with radiotherapy due to the complex geometry of the HN region, and close proximity of organ-at-risks (OARs) such as the spinal cord, brainstem, parotid glands and optic structures.

HN is a clear example of a patient cohort known to benefit from ART, and therefore the focus of clinical research. Several studies show that HN patient's anatomy can change considerably during the course of the treatment [43, 59–61], and that this results in dosimetric changes from the original plan [46, 62–68]. It is clear that some patients require at least one replan [46, 56, 69–73], but it is not clear which benefit the most from ART and when is the right time for intervention.

HN patient data was used in the initial phase of this project to explore and optimise NiftyReg DIR (this chapter). Follow-up studies on the clinical applications were investigated in chapters 3, 4, 5 and 7.

2.3.2 The clinical problem: lung

Lung cancer is the leading cause of cancer-related death in the USA and worldwide. There are two major types of lung cancers: small cell lung cancer and the more predominant non small cell lung cancer. In 2015, the National Cancer Institute estimated 220000 new cases and 158000 deaths for both men and women in the United States, with a 5-year survival rate of only 17.4% [74].

The overall low survival rate of lung cancer led to new treatments that aimed to improve local and locoregional controls, such as radiation dose escalation [75, 76]. However, when delivering higher dose with curative intent, both the dose delivered and the expected tissue toxicity must be considered, with lung itself often being the most dose-limiting organ at risk [77–79]. To limit risks of radiation-induced injuries to normal tissues such as lungs, heart, oesophagus and cord, higher tumour doses are not always achievable [80].

Intra-fractional changes are the most researched aspect of lung tumours [81–83]. In lung radiotherapy several planning and delivery strategies have been developed to consider and reduce the effects of motion. Four-dimensional (4D)-CT is acquired with the patient free-breathing, and the planning is conducted on an average CT resulting from all the phases of the 4D-CT. The gross tumour volume (GTV) then is delineated on all the frames and the planning GTV consists of the union of all motion phases, which after expanded by the margin for suspected microscopic disease forms the internal clinical tumour volume (iCTV). In some cases, the intensity of the target in the average-CT is replaced by the maximum intensity projection (MIP), i.e., the voxel of greatest value from all corresponding voxels over the respiratory cycle. During delivery several approaches are suggested to mitigate or compensate for the effects of motion, such as breath-hold, jet-ventilation, gating, and beam tracking. However other inter-fractional changes during the course of radiotherapy may also affect the dose delivered to target and healthy tissues [84, 85]. These factors include, but are not limited to, changes in tumour size and position, alterations in tissue anatomy, variations in respiratory patterns, and fluctuations in patient weight [86].

The lung cohort will be the focus of interest of chapter 6.

2.4 Initial studies: optimisation of NiftyReg

Before clinical use, any DIR algorithm should be validated within the context of the specifications of the desired implementation, the clinical environment (modalities, image quality, sites) and user-defined parameters [47]. This section describes the studies performed at the beginning of this project, to familiarise with the concepts and software tools. It also allowed to understand the problems associated with the evaluation and validation of DIR.

In this preliminary study a large range of NiftyReg parameters were investigated on two HN datasets to find a set of promising parameters of DIR to use in CT-to-CBCT studies. This process generates a deformed CT (dCT), that matches the geometry of the CBCT. Ideally, DIR should be used automatically in clinical applications. Therefore a set of parameters was investigated such that it gave good results (but not necessary optimal) for each dataset by minimising the computation time and keeping acceptable values for the similarity measures. Since high values of similarity measures do not necessarily mean a better registration the analysis is aided with visual assessment of the registered images and corresponding deformed grids. Additionally, analysing the deformation itself besides the deformed image is very important to assess if the transformations are well regularised (Figure 2.5). Understanding the underlying performance of the algorithm is fundamental before translating DIR to clinical applications.

2.4.1 Choice of registration parameters and algorithms

Different user-defined parameters of NiftyReg were extensively studied in this section (see Table 2.2). This included the effect of the initialisation (i.e., initial rigid alignment), masking (i.e., ignore regions of the image during optimisation), CP grid size, penalty terms, similarity measures and implementations of the B-spline algorithm.

Rigid only transformations were applied to describe the global alignment between CT and CBCT, as it is an intra-subject registration. Mask usage in the rigid registration improved the global alignment results and reduced significantly the computation time. Ignoring the last level of resolution has no visual effect if masks are used, which reduces the computation time even further. The rigid registrations took approximately 1 minute to compute. The rigid registration algorithm has other parameters that can be tuned. When the datasets are well aligned to begin with, and no large rotations are necessary, the remaining default parameters are adequate. A rigid-only registration could not fully capture all the changes that occur in the HN region (Figure 2.6); thus, DIR was necessary. Regarding the deformable registration,

1. Good results were achieved when using a control point spacing (CPS) between 5 and 10 voxels. For lower values it is difficult to sufficiently constrain the registration and the algorithm loses the ability to capture bigger deformations. It is appropriate to

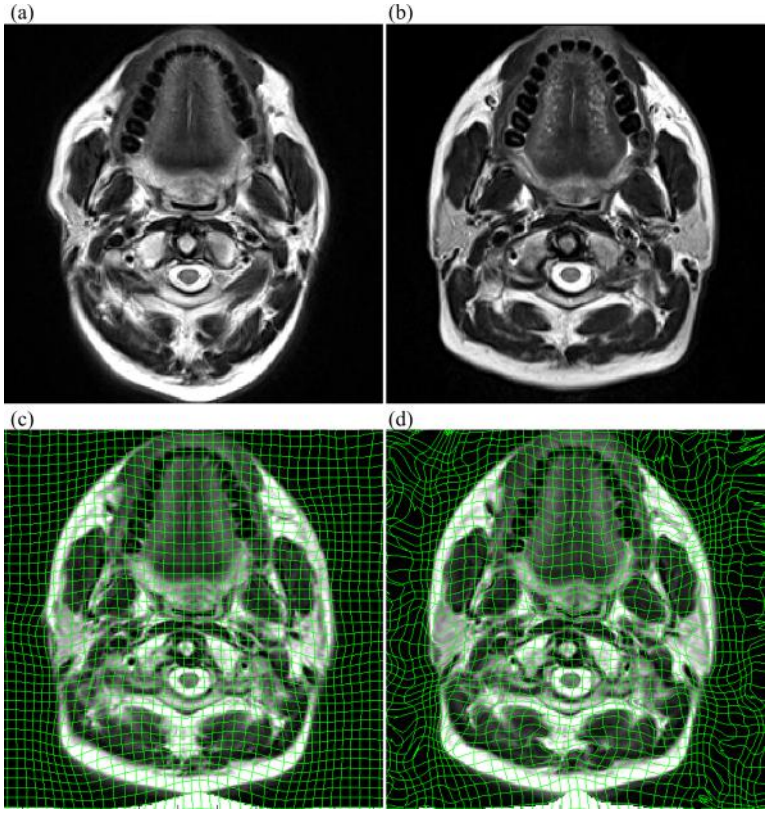

Figure 2.5: *Importance of well constrained registrations. Example of (a) source and (b) target images, and a (c) under-constrained and (d) well-constrained transformation. The under-constrained registration may be more similar to the target image, but physically implausible deformations are occurring. For applications where the underlying deformation is important such deformations are not adequate. Courtesy of Dr Jamie McClelland.*

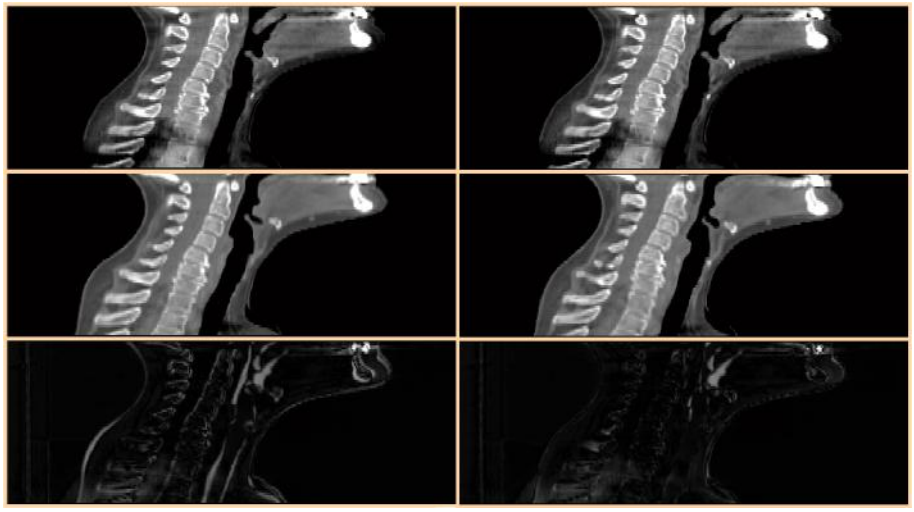

Figure 2.6: *Saggital slices of the CBCT, registered image and difference image between the two. The first column refers to a rigid-only transformation, and the second to the deformable transformation. Grey areas show where the CBCT and registered image disagree. Even though the anatomy of the head and neck is conventionally considered rigid, a rigid-only registration cannot fully capture all the changes. A rigid registration shows considerable disagreement in the bone and external contours alignment. Using deformable registration the matching is improved. Near the throat there are still discrepancies due to swallowing.*

first find the value of CPS that captures the deformation, and then tune the weight of the penalty terms accordingly. Using voxels instead of mm to define the CPS made the automated process more foolproof, i.e., if the resolution of datasets varied between patients without the user realising, it could result in using more CPs than voxels, leading to over-parametrisation of the registration.

2. The similarity measure used for CT-CBCT registration was the NMI (equation 2.6). Tests conducted with SSD as similarity measure showed that it did not perform well due to the intensity differences between CT and CBCT images. On a later stage of this project, LNCC became available as well, and proved to be more appropriate for monomodal and quasi-monomodal registrations. This similarity measure was used in chapters 5 and 7.

3. The similarity measure used may also have specific parameters to be tuned. In the case of NMI, the number of bins used in the joint histogram calculations is used and was found to affect the overall results. Unexpectedly, increasing the number of bins did not seem to increase the computation time but degraded the final registration. It is possible that a higher binning value makes the code more sensitive to noise in the images. Also, CBCT intensity values for the same type of tissue may vary in different areas of the image; thus, smaller binning intervals may lead to the same tissue being separated in different bins on different zones of the image. Considering NMI values and visual assessment, a binning between 32 and 128 was acceptable in all tests done. When using LNCC as measure of similarity, the user can define the standard deviation of the Gaussian kernel used to assess the pixel's neighbourhood. Studies performed for the optimisation of DIR for chapters 5 and 7 suggest a value within the interval [8, 20].

4. Values of BE within the interval [0.01, 0.10] appeared to produce acceptable results. For these two datasets the best visual results were produced by a narrower range ([0.02, 0.06]). Low values of BE resulted in higher NMI similarity values (Figure 2.7), but visually the alignment could be incorrect .

5. In general, using JL only causes dramatic visual changes for low values of BE, where unrealistic deformations such as folding are more likely to occur. The introduction of this parameter smoothed the effect that other parameters changes had. Thus values within [0.01, 0.10] seemed like a good compromise between constrain of the transformation and stopping folding from occurring. Besides contributing to the cost function, the use of JL prompts a folding correction scheme to the final DVF.

6. A thresholding may be advisable to remove "padding" values (voxels with intensities less than -1000 HU) and to deal with high-intensity artefacts. The thresholding conducted showed no improvement in image alignment and computation time. However, since the effect of threshold is similar to the effect of increasing the binning of the joint histogram, for thresholding to have a positive effect the choice of

binning had to be adjusted properly. Further testing showed that using a thresholding of [-1000 2000] combined with a smaller binning of 32 seemed to improve the external contours and bony anatomy.

7. Reducing the maximum number of iterations reduces the computation time by forcing the algorithm to finish before it reaches a convergence value. Overall was a good compromise to use a maximum of 1000 iterations at, which limited the maximum computation time to approximately 5 minutes on CT-CBCT datasets.

8. NiftyReg implements a multi-resolution approach, and the number of levels can be defined by the user. Since the datasets were well aligned to begin with, changing the number of levels of resolution did not bring great changes in NMI, computation time and visual quality. Three resolution levels were appropriate for the datasets available.

9. The choice of rigid registration parameters does have an impact in the following deformable registration. In general, initialising the deformable registration with a better rigid alignment reduces the time spent to reach convergence (Figure 2.8). The final results were visually similar while overall the NMI value was slightly improved.

10. Masking the reference image in the deformable transformation reduced the computation time to 1 minute, but made the registrations more sensitive to other parameters.

11. The symmetric inverse-consistent algorithms were still in development during the initial stage of this project. Both this and the stationary velocity fields implementations were further studied from chapter 4 onward. These two implementations were more robust to unrealistic deformations which allowed to reduce the weight of the regularisation terms in comparison to the default asymmetric algorithm.

2.4.2 Image pre-processing to improve registration quality

As a follow-up of the preliminary study, the impact of image pre-processing (to reduce noise and enhance contrast) on the registration quality was investigated. Other authors suggested that noise reduction and image contrast enhancement could improve the registration accuracy and convergence speed [55, 87, 88]. With SSD as similarity measurement one expects pre-processing to contribute for a better registration. However with NMI that may no longer be the case, as the registration is less sensitive to the actual intensity values.

Lu et al. suggested the application of an "edge-preserving filter" to the CBCT image prior to the deformable registration, to take in consideration the differences between the modalities and to minimize their effect in the resulting transformation [55]. Conventional denoising techniques, such as Gaussian smoothing, tend to blur the sharp boundaries

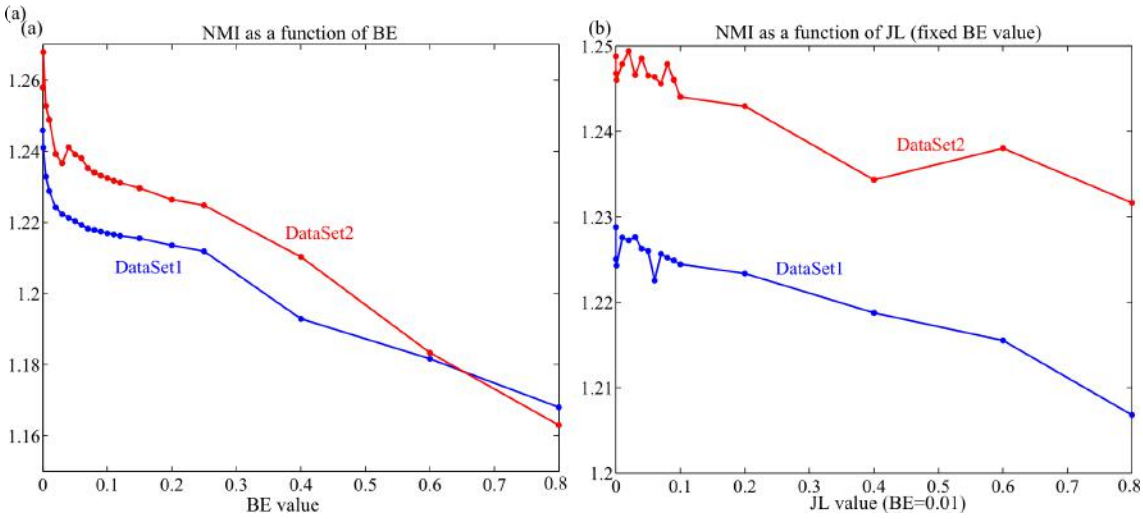

Figure 2.7: *Decreasing of the normalized mutual information (NMI) value with increasing value of the penalty terms: bending energy (BE) and logarithm of the determinant of the Jacobian (JL).*

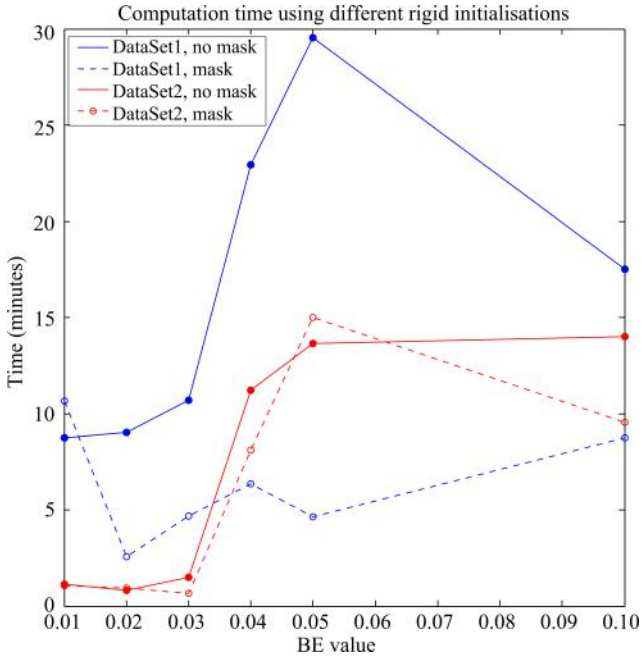

Figure 2.8: *Computation time as a function of bending energy penalty term (BE) using different initial rigid alignments. Mask stands for initialisation where a mask was used during the rigid registration, which resulted in better initial alignment.*

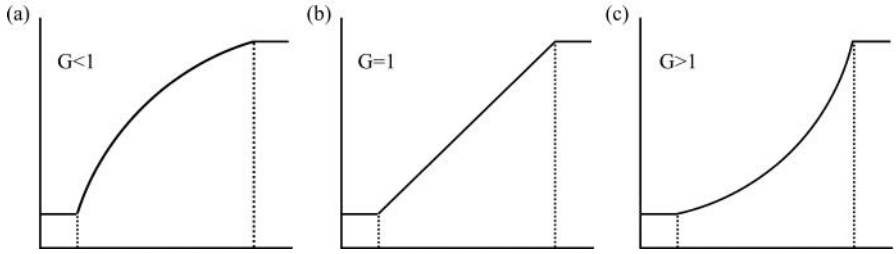

Figure 2.9: *Transformation applied by gamma correction for different gamma values.*

within the image. Such filter is not appropriate for DIR as loosing information at boundaries will degrade the registration. An anisotropic diffusion filter however attempts to smooth very little at sharp boundaries and smooth more when no boundaries are present. This is achieved by modifying the classic diffusion equation into the following form [89]:

$$\frac{\partial I(\mathbf{x},t)}{\partial t} = \mathbf{div}[g(\|\nabla I\|)\nabla I] \tag{2.12}$$

where ∇I is the image gradient and $\|\nabla I\|$ is the gradient magnitude. $g(\|\nabla I\|)$ is the "edge-stopping" function. In the limit where $g(\|\nabla I\|)$ is a constant, equation (2.12), becomes the classic diffusion equation (and a Gaussian smoothing is applied). A possible option for the "edge-stopping" function is as follows [89]:

$$g(\|\nabla I\|) = e^{-\frac{\|\nabla I\|^2}{2k^2}} \tag{2.13}$$

The diffusion is slowed or even stopped in the edges, where the image gradient is large ($\lim_{x\to\infty} g(x) = 0$). The parameter k controls the sensitivity of the smoothing process. If a small value is used, large amplitude noise may be preserved, while if a large value is used image detail is lost.

In gamma correction the intensity values are mapped to a specified range, so that the intensities can be better distributed on the histogram. Depending on the value of gamma (G) the mapping can be linear or non-linear. G can take any value in the interval [0, ∞[. If G=1 the mapping is a linear transformation. However if G <1 the mapping is weighted toward higher (brighter) output values, and if G>1 the mapping is weighted toward lower (darker) output values (Figure 2.9).

The effects of the following pre-processing techniques in the DIR were investigated:

- Edge-preserving smoothing, using a 3D implementation based on Gerig et al. [90];
- Intensity adjustment, using MATLAB (MathWorks, Natick, MA, USA) Image Enhancement toolbox.

The following values were tested: k={0, 50, 100, 150} and G={1.0, 1.5, 2.0}. G>1 make a better use of the dark region of the spectrum to better differentiate soft tissues. If G>2.0 the image would get too dark and information from lower intensities was lost. The range G=[1, 2] seems to have a good compromise between contrast and tissue information.

Visual assessment and variation of NMI for different pre-processing processes showed that:

- **Edge-preserving filtering**: Visually there were not relevant changes between using pre-processing or not. Most of the differences were within the soft tissue region, and therefore harder to evaluate with visual inspection. In some areas the alignment seemed to be better, while in others the opposite occurs. Smaller values of k also seem to be a better option, since some deformations start to appear for k=150. The changes in NMI using edge-filtering were negligible.

- **Gamma correction**: While a value of G=2.0 seems to degrade the registration, a smaller value as G=1.5 seems to improve the registration results, particularly in the bony alignment and external contours.

Gamma correction was a viable option to improve the registrations by pre-processing the reference image, while edge-preserving filtering had an impact on the final registration but it was hard to quantify if there was a benefit or not using visual inspection. The pre-processing methods studied did not seem to improve the computation time.

2.5 Optimisation of NiftyReg for CT to cone-beam CT deformable image registration

In section 2.4 the user-defined parameters available in NiftyReg were explored, and a large range of parameters that provided suitable visual matching of the registrations was found. NiftyReg DIR was found to be fairly robust as small changes in the parameters did not cause dramatic changes in the registration results, particularly when the Jacobian penalty term was used.

For rigid registrations the best results were achieved using a mask and ignoring the last level of resolution. The optimal parameters found for the DIR were: BE=[0.02, 0.06], CPS=[0.01, 0.10], 3 levels of resolution, a maximum number of 1000 iterations, a binning of 32 for the joint histogram calculation, thresholding the intensities to the range [-1000, 2000] and a CPS between 5 and 10 voxels. Such registrations can took up to a maximum of 5 minutes to finish; using a mask this time was reduced to less than 1 minute. Pre-processing the CBCT image with non-linear intensity adjustment could also improve the registrations.

In this section the aim was to narrow the large range to a single set of parameters that performs well over all the datasets available, and identify strategies to quantitatively evaluate and validate DIR results.

2.5.1 Methods and Materials

2.5.1.1 Patients data acquisition

Data from five HN patients that were identified as potential candidates for replanning and referred for dosimetric assessment at UCLH was used in this study. The referral happened when the spinal canal or brainstem was found outside their respective planning organ at risk volume, their external contour had decreased more than 5 mm and/or the immobilisation mask was no longer effective.

This cohort had therefore large anatomical variations, and was a challenging dataset for DIR which contained only patients that would benefit from CBCT and DIR-based ART.

All patients underwent IMRT with a planned dose of 65 Gy delivered in 30 daily fractions. Patient positioning was assured by appropriate head-rest and a personalised HN and shoulder mask.

The imaging protocol consisted of a pCT (GE Widebore 16 slice system, GE Healthcare, Little Chalfont, UK) with contrast injection, and weekly CBCTs (On-board imaging v1.4, Varian Medical Systems, Palo Alto, CA, USA) acquired in treatment position. The CBCTs were acquired in half-fan mode, full rotation, 110 kVp, 20 mA, 20 ms, with a maximum FoV of 45 cm in diameter and 16 cm in length. Imaging resolution was 0.977×0.977×2.5 mm^3 and 0.879×0.879×2 mm^3 for the CT and CBCT scans, respectively.

2.5.1.2 Registration settings

Four different parameter settings and/or were compared: (i) different BE (3% and 5%), (ii) different CPS (5 and 10 voxels), (iii) pre-processing of the CBCT and, (iv) masking, giving a total of 16 registrations per dataset. The remaining parameters were kept constant, and matched those described in section 2.5.

2.5.1.3 Contours comparison

Registrations were compared qualitatively, by visual inspection, and quantitatively, by computation time and similarity of warped structures with a gold-standard. The deformation field resulting from the registration can be used to map points and regions of interest delineated in the pCT dataset to the CBCT dataset. Structures delineated in the pCT were warped and compared with contours manually drawn by the same physician on the CBCT. Due to the noise and low contrast inherent to CBCT imaging, it is difficult to define points or delineate structures with confidence and consistency. The features were chosen to be structures that could be unequivocally identified in both scans and that gave an indication of how well the registration accounts for patient positioning errors and

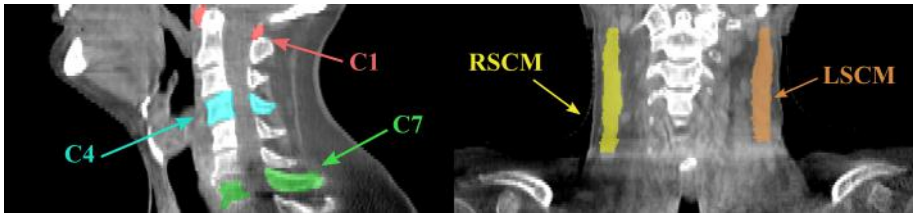

Figure 2.10: *Structure set manually delineated on the CT and CBCT of each patient.*

weight loss. Vertebrae C1, C4 and C7 were considered as bone landmarks for different regions, while external body and sternocleidomastoid muscles, right sternocleidomastoid muscle (RSCM) and left sternocleidomastoid muscle (LSCM), were chosen as surrogates for soft tissue (Figure 2.10). The structure set was delineated in all datasets by Syed Moinuddin (Department of Radiotherapy, UCLH).

Vertebrae C1, C4 and C7 were used because they are only subject to rigid motion (i.e., they do not deform), and cover the length of the cervical spinal canal. External body contour, RSCM and LSCM were used as soft tissue structures. The two muscles are adjacent to the region that contains the neck lymph nodes. Deformation between scans may affect their shape and position, and therefore nodal dose [91].

Considering that A and B are the set of voxels that define the volumes of the manual and deformed features while $\tilde{A}$ and $\tilde{B}$ define the corresponding surfaces, four metrics were used to describe the similarity between the features:

- Dice similarity coefficient (DSC), which describes the overlapping ratio between two volumes of interest.

$$\mathrm{DSC} = \frac{2|A \cap B|}{|A| + |B|} \tag{2.14}$$

- Overlap index (OI) can be used as an additional measure of volume overlap to consider one as a gold-standard, since DSC does not allow preference between two volumes [51]. The manual contours are the gold-standard (A):

$$\mathrm{OI} = \frac{|A \cap B|}{|A|} \tag{2.15}$$

- Distribution of Euclidean distances between surfaces' points, also known as the distance transformation (DT) [54].

$$\mathrm{DT}(\mathbf{a}) = \min(\|\mathbf{a} - \mathbf{b}\|), \quad \mathbf{a} \in \tilde{A}, \quad \forall \mathbf{b} \in \tilde{B} \tag{2.16}$$

  The algorithm written to compute DT is signed (i.e., values can be positive or negative depending if the surface is surrounding or within the gold-standard volume). Since DT is a distribution and not a single value, several statistical quantities can be derived from the distribution to facilitate the analysis of the results. The statistics are computed bi-directionally.

- Center-of-mass (centroid) position error (CoM).

$$\text{CoM} = \|\mathbf{a}_{\text{CoM}} - \mathbf{b}_{\text{CoM}}\| \tag{2.17}$$

The metrics presented provide complementary information about the overlap between volumes (DSC and OI), closeness between the surfaces (DT) and spatial positioning of the features (CoM). In this study DT results were summarised in terms of the fraction of the DT distribution larger than a tolerance value of 2 mm ($|\text{DT}|_{2\text{mm}}$).

2.5.1.4 Dosimetric analysis

Dose comparisons were performed on three patients to assess whether reduced accuracy but faster registrations had a noticeable effect when performing dose calculations. Dose calculations for the IMRT plans that the patients were treated with Eclipse (Varian Medical Systems, Palo Alto, CA, USA) External Beam TPS analytical isotropic algorithm (AAA), using the highest available resolution (1 mm).

The comparison of dose distributions is conventionally performed using a number of different methods:

- Dose-volume histograms (DVHs): DVHs summarise the dose distribution information within a volume of interest (VOI). The more commonly used DVHs in practice are the cumulative type, which are plots of volume receiving a dose greater than (or equal to) a given dose, against dose [18].
- Dose-difference (DD): absolute difference, voxel-by-voxel, between two dose distributions [92]. This quantity is very sensitive in high dose gradient regions.
- Distance to agreement (DTA): distance between a dose point in a gold-standard dose and the nearest point in the measured dose distribution containing the same dose value. This is equivalent to determining the shortest distance between the dose point at the reference and the isodose surface of the evaluated dose distribution [93]. It is more suitable than DD for high dose gradient regions, but it is overly sensitive in low-dose gradient regions.
- Gamma-index: DD and DTA alone can be insufficient, and the two methods actually complement each other. The gamma analysis method compares a reference (D_{ref}) and calculated (D_{cal}) dose distributions using acceptance criteria [94]. It combines two important dose comparison criteria: DTA (Δd_M) and DD (ΔD_M). The gamma index (γ) at each point of the dose distribution is given by:

$$\gamma(\mathbf{r}_{\text{ref}}) = \min\{\Gamma(\mathbf{r}_{\text{cal}}, \mathbf{r}_{\text{ref}})\} \forall \{\mathbf{r}_{\text{cal}}\} \tag{2.18}$$

where

$$\Gamma(\mathbf{r}_{\text{cal}}, \mathbf{r}_{\text{ref}}) = \sqrt{\frac{r^2(\mathbf{r}_{\text{cal}}, \mathbf{r}_{\text{ref}})}{\Delta d_M^2} + \frac{\delta^2(\mathbf{r}_{\text{cal}}, \mathbf{r}_{\text{ref}})}{\Delta D_M^2}} \tag{2.19}$$

with

$$r(\mathbf{r}_{\text{cal}}, \mathbf{r}_{\text{ref}}) = |\mathbf{r}_{\text{cal}} - \mathbf{r}_{\text{ref}}| \tag{2.20}$$

and

$$\delta(\mathbf{r}_{\text{cal}}, \mathbf{r}_{\text{ref}}) = D_{\text{cal}}(\mathbf{r}_{\text{cal}}) - D_{\text{ref}}(\mathbf{r}_{\text{ref}}) \tag{2.21}$$

The pass-fail criteria are: $\gamma(\mathbf{r}_{\text{ref}}) \leq 1$ the calculation passes, $\gamma(\mathbf{r}_{\text{ref}}) > 1$ the calculation fails. The criterion used was Δd_M=2 mm and ΔD_M=2%.

The dose distributions were compared considering DDs and gamma analysis. In terms of DDs the test absolute mean value ($|\text{DD}|_{\text{mean}}$) and pass-percentage ($\text{DD}_{\text{2\%-pp}}$) , i.e., the percentage of pixels whose DD was inferior to 2% of the prescribed dose (pD). For the gamma-index the test pass-percentage, i.e., the percentage of pixels whose gamma-index was inferior 1, was calculated. 2%pD and 2 mm were defined as tolerance criteria for DDs and distance in accordance with UCLH's internal clinical standards for the comparison of dose distributions.

The tools needed for the geometric and dosimetric evaluations were implemented; freely available code was used and modified when relevant [95, 96].

2.5.2 Results

The registrations were grouped by varying one of the four parameter settings and/or approaches: (i) different BE (3% and 5%), (ii) different CPS (5 and 10 voxels), (iii) pre-processing of the CBCT and, (iv) masking, giving a total of 16 registrations per dataset. For example, all the registrations were separated in two groups: those with a bending energy weight of 3% and those of 5%. The four approaches were compared in terms of their effect on the different quantities averaged over all structures and datasets (Table 2.3), and for individual structures and datasets:

- Different BE: a lower bending energy gave better results in 66% of the cases. It had a general good impact in all quantities, particularly for soft tissue structures. Lower BE implies more freedom to perform extreme contractions/expansions, so it was easier to capture more extensive tissue deformations.
- Different CPS: smaller CPS gave better contour statistics on 59% of the cases. It particularly improved statistics for soft tissue surrogates.
- Choice of preprocessing the reference image with non-linear intensity adjustment: this pre-processing improved contour statistics on 70% of the cases. All statistics but OI were improved, particularly for bones and body outline. This reduction in OI might be related with reducing the overestimation of the volume of the structures, as schematically explained in Figure 2.11.

Table 2.3: *Mean ± standard deviation of dice similarity coefficient (DSC), overlap index (OI), distance transform (DT) and centroid position error (CoM) grouped by varying parameter.*

Parameter	Value	DSC	OI	$\|DT\|_{2mm}$ (%)	CoM (mm)
Bending Energy weight	3%	0.843±0.012	0.867±0.015	8±3	1.4±0.2
(BE)	5%	0.837±0.013	0.861±0.016	8±4	1.4±0.3
Control Point Spacing	5	0.851±0.006	0.877±0.014	7±3	1.5±0.3
(CPS)	10	0.828±0.005	0.852±0.011	9±4	1.4±0.2
Masking the reference	Yes	0.839±0.014	0.862±0.013	8±3	1.4±0.3
image	No	0.841±0.013	0.867±0.015	8±3	1.4±0.2
Pre-processing the	Yes	0.842±0.014	0.864±0.012	8±3	1.4±0.2
reference image	No	0.837±0.012	0.866±0.016	9±4	1.5±0.3

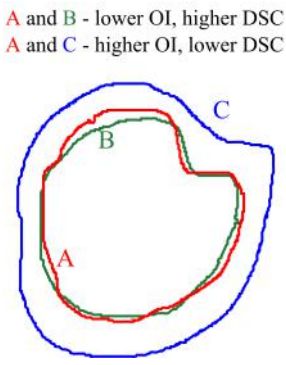

Figure 2.11: *Advantage of using dice similarity coefficient (DSC) over overlap index (OI). Overestimating the volume of a structure results in higher OI, which may not be indicative of a better matching.*

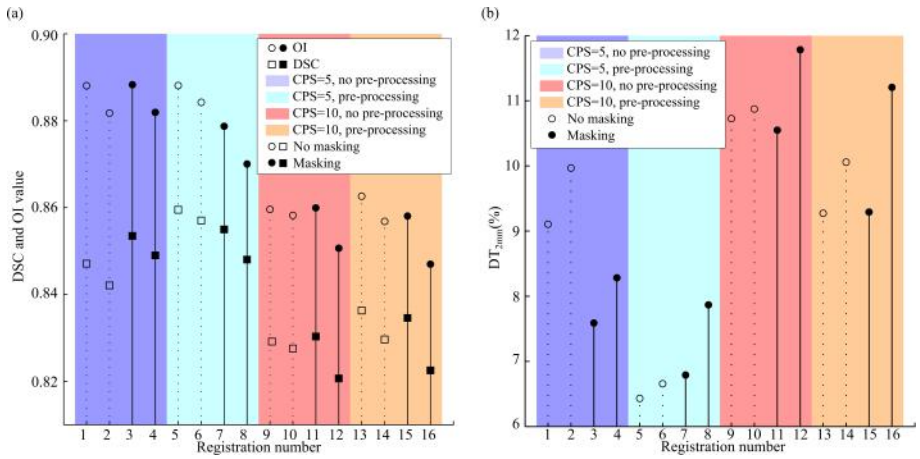

Figure 2.12: *Variation of a) DSC (squares) and OI (circles) and b) $|DT|_{2mm}$ for different combinations of parameters: BE=3% (#1,3,5,7,9,11,13,15) and BE=5% (#2,4,6,8,10,12,14,16); CPS=5 (#1-8) and CP=10 (#9-16); mask (#3,4,7,8,11,12,15,16) and no mask (#1,2,5,6,9,10,13,14); pre-processing (#5-8,13-16) and no-preprocessing (#1-4,9-12). Smaller control point spacing translated in larger values of DSC and OI. Pre-processing the CBCT improved DSC and $|DT|_{2mm}$ values.*

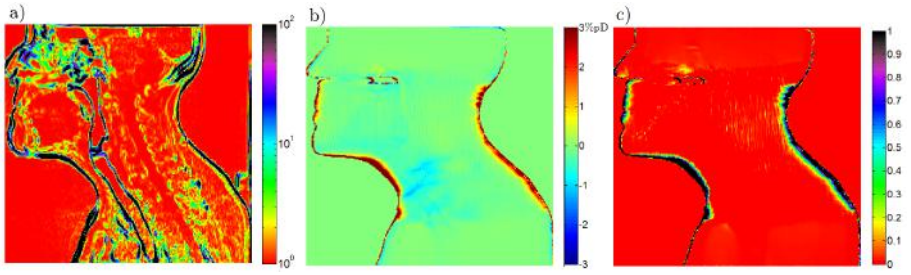

Figure 2.13: *For the two registration approaches: a) difference in pixel intensity, b) dose-differences and c) gamma-index. The similarity between the images is high, with most of the differences smaller than 10 HU. The regions where the differences in dose were higher occurred outside the CBCT field-of-view, in the skin, immobilisation mask and airways, and these had little effect on the dosimetry of the OARs.*

- Choice of masking: masking gives better contour statistics only on 44% of all the cases. Like visually assessed, masking has an unpredictable effect on the registrations accuracy.

These findings suggested that combining pre-processing, CPS=5 and BE=3% resulted in better registrations. The benefits of masking were not obvious, since it appeared to slightly degrade the registrations. The differences in CoM were negligible in all cases, as the variations were a fraction of the dimensions of the pixel.

This assessment considers that the effect of a parameter on the registration is independent of the other parameters, which is not the case. To confirm that the chosen parameters did stand out compared with the rest, the mean DSC, OI, $|DT|_{2mm}$ and CoM were analysed for each registration separately. Figure 2.12 shows how the registrations in the light blue region (CPS=5 and pre-processing) with odd number (BE=3%) gave the best results in general.

Dose distributions were calculated by applying the same IMRT plan to the registration results that used a mask, and those that did not. The goal was to assess if accuracy for speed trade-off had a big impact in the dose calculated. The two dose distributions obtained were similar. DD was smaller than 2%pD on 96.2±0.4% of the voxels, with a mean value of 0.38±0.01%pD. The comparison passes a gamma-test on 99.4±0.1% of the voxels. Figure 2.13 gives a qualitative overview of the results obtained.

2.5.3 Discussion

Properly choosing the registration parameters is one of the most important steps in using DIR on clinical data as the parameters determine the actual deformations produced. The parameters cannot be tuned for each individual case, but should be tuned for different sites, imaging modalities, image size and quality, etc. Often default parameters which work well on a wide range of different images will be sub-optimum for a particular registration task.

Lower BE allowed more flexibility for the algorithm to perform larger contractions and expansions, however if its weight was too small the registrations were under-constrained and unrealistic deformations were likely to occur. The JL encouraged the registrations to be folding-free and ensured one-to-one mapping.

Non-linear intensity adjustment of the CBCT image increased the contrast between structures, sharpening the anatomy boundaries. This process helped the DIR distinguishing different structures. This was reflected on higher DSC and closer matching of warped structures' surfaces (better $|DT|_{2mm}$ values), at the price of lower OI values (Figure 2.12).

Masking had an unpredictable effect on the registrations accuracy, depending on the particular dataset. The reduction of computation time in one order of magnitude is good reasoning for always masking the registrations, and the differences in accuracy had a negligible effect on the dose calculations (Figure 2.13).

2.6 Conclusions

In this chapter, the preliminary investigations into the major points of interest of this thesis were performed. The concepts of DIR and ART were introduced, together with the clinical problems. An in-house DIR algorithm (NiftyReg) was tested and investigated in detail to achieve excellent performance at CT-to-CBCT registrations in the HN site, for images obtained in our clinic. The optimal parameters found for the deformable registration were: BE of 3%, JL of 1%, 3 levels of resolution, a maximum number of 1000 iterations, a binning of 32 for the joint histogram calculation, thresholding the intensities to the range [-1000 2000], CPS of 5 voxels, non-linear intensity adjustment of the CBCT and masking. The understanding achieved of the DIR software provided the expertise

necessary to tune NiftyReg registrations for other sites and imaging modalities, as will be the case in later chapters of this thesis. In such cases, the full detailed optimisation process was omitted of this thesis since a similar protocol was followed. The work performed in this section was also fundamental to identify strategies for the geometric and dosimetric evaluation of DIR. Finally, the tools here described for the evaluations and data communication with TPS were used multiple times throughout the remaining of this thesis.

Chapter 3

Cone-beam CT and deformable image registration for "dose of the day" calculations

> One never notices what has been done; one can only see what remains to be done.
>
> Marie Curie

NiftyReg was optimised for CT-to-CBCT registrations in HN patients in chapter 2. ART workflows require an assessment of the dose delivered to the "anatomy of the day", i.e., the "dose of the day"; in this chapter the use of CT-to-CBCT for this clinical application is explored.

The work in this chapter resulted in the following outputs:

- C. Veiga, J. McClelland, S. Moinuddin, A. Lourenço, K. Ricketts, J. Annkah, M. Modat, S. Ourselin, D. D'Souza, and G. Royle, "Toward adaptive radiotherapy for head and neck patients: feasibility study on using CT-to-CBCT deformable registration for "dose of the day" calculations," Med. Phys. 41 031703 (2014).
- C. Veiga, J. McClelland, S. Moinuddin, K. Ricketts, D. D' Souza, and G. Royle, "Calculation of the dose of the day using an in-house validated deformable registration algorithm," Radiother. Oncol. 106(S2), S478 (2013). (Geneva, Switzerland, 2013).

3.1 Rationale

To calculate the "dose of the day" and assess if the current plan is still acceptable, an image of the patient in treatment position with structures of interest delineated is

necessary, and DIR can provide a solution to answer both those needs. CT-to-CBCT DIR can be used to map the HU information from the planning to the daily geometry, and therefore assess the actual dose delivered in each treatment fraction. Two other common approaches suggested in the literature to calculate the "dose of the day" are based on image guidance with CT imaging [69], and direct dose calculations on the CBCT, using pixel correction techniques [97, 98] or relative electron density (RED) calibration [9, 99]. The first increases the dose given to the patient in the image guidance protocols and requires an in-room CT scanner, which is not available at UCLH. The second is more limited by the inherent properties of CBCT imaging, such as proneness to motion artefacts, increased noise, reduced contrast and limited FoV. Treatment planning also requires delineation of structures of interest, which can be challenging in a CBCT scan.

DIR validation is challenging due to the lack of gold-standards in clinical and non-clinical settings [100]. While there is a wide variety of studies assessing the quality of CT-to-CT deformable registration with patient data [51, 54], for CT-to-CBCT the studies are scarcer and usually focused on the deformation properties rather than dosimetry [10, 88, 101, 102]. In this chapter image inspection, feature-based evaluation and comparison of dose distributions were used to assess the suitability of DIR for the clinical application of dose calculations in HN patients.

3.2 Methods and Materials

3.2.1 Patient data acquisition

Retrospective data from five HN patients treated at UCLH and referred for possible replan was used in this study. The same cohort that was described in Section 2.5.1.1 was used. Table 3.1 provides additional details of the patients included.

Replanning referral occurred when the CBCT offline review study found the spinal canal or brainstem outside their respective planning organ at risk volume, the external contour decreased more than 5 mm and/or if the immobilisation mask was no longer effective. The treatment isocenter is usually set to bony anatomy on the identifiable vertebrae, and does not represent any normalisation point or high dose region. The CBCT is aligned to the pCT following a standardised online image-guidance protocol for isocenter alignment based on manual rigid registration to the cervical spinal vertebrae. By rigidly aligning the pCT with the CBCT and defining the new body external contour on the pCT based on the CBCT external, target coverage and possible overdosage to OARs were verified and the decision to replan taken. Four of these patients were replanned midway. A replan/rescan CT (rCT)[1] was acquired in the same scanner as the pCT (with

[1]The terms replan and rescan CT are often used for the same purpose, but have a subtle difference in definition. Rescan means that the same positioning as planning was reproduced; in a replan CT the positioning may change. Therefore, in this chapter the most correct term is replan CT.

Table 3.1: *Characteristics of the patients included in the study.*

Pt no	Age (y)	Gender	Tumour site	TNM classification	Replan (Y/N)	ΔW[a] (%)	ΔV_{ext}[b] (%)	L_{PTV}[c] (mm)	V_{PTV}[d] (%)
1	64	F	Oropharynx	T3N1M0	N	N/A	-8.5	+15	6.5
2	61	M	Larynx	T3N1M0	Y	+0.4	-3.5	+8	0.5
3	73	F	Base of the tongue	T4N2cM0	Y	-1.7	-4.7	0	0.0
4	60	M	Larynx	T3N0M0	Y	-4.4	-12.4	+14	3.0
5	64	M	Pharyngeal wall	T4N2cM0	Y	-11.7	-8.6	-26	9.4

[a] ΔW = relative weight variation at plan evaluation

[b] ΔV_{ext} = relative external volume variation at plan evaluation, in the region imaged by the CBCT;

[c] L_{PTV} = length of the target volume outside the CBCT in the superior/inferior direction, at replan referral;

[d] V_{PTV} = target volume fraction not imaged by the CBCT, at replan referral

contrast) and a new plan built from scratch; a new immobilisation mask was necessary but the previous positioning was reproduced as close as possible. Typically the last two weeks of treatment were completed with the new plan.

3.2.2 Image registration settings

The registrations were defined by a set of parameters optimised to suit the datasets being registered, discussed in chapter 2. The standard, unidirectional and asymmetric implementation on the GPU of NiftyReg was used. Two regularisation terms were used: BE, which encourages a smoothly varying deformation field [21] and JL, which penalises large volume changes [36]. A folding correction scheme is applied every iteration and to the final deformation field. When folding occurs the correction scheme updates the CP coefficients in the vicinity of folded voxels to try and produce a folding-free transformation [103]. The folding correction scheme ensures the invertibility of the DVF. NMI was preferred as similarity measure over other popular measures, such as the SSD, since it not only handles the non-linear relationship between CT and CBCT intensities but also the local variations of intensity characteristic of CBCT imaging.

The patient images and structures were exported from the TPS in DICOM format to a standalone registration workstation. The registration workstation had an Intel Xeon CPU E25606 (2.13GHz, 12GB RAM) with a NVIDIA Tesla C2070 GPU card (14 multiprocessors, 6GB dedicated memory). A rigid registration was first applied in order to estimate the global alignment between the pCT and the CBCT. The obtained transformation was then used to initialise the deformable registration. The deformable registrations using images at full resolution ran in approximately 1 minute. The output DVF was used to propagate

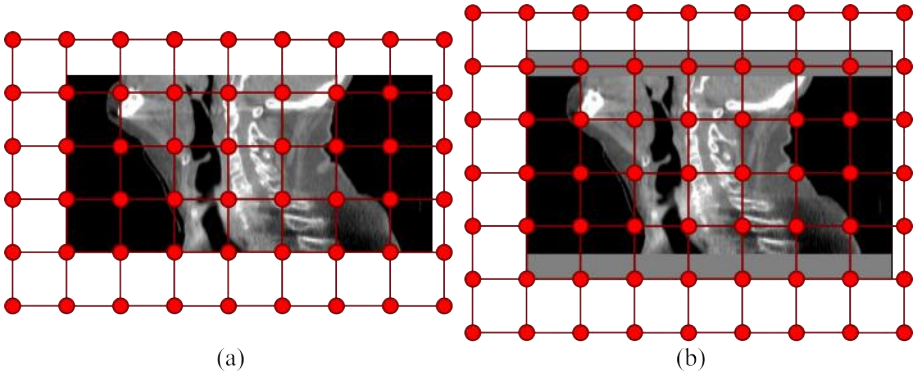

Figure 3.1: *Standard B-spline control point grid covering (a) the CBCT volume only and (b) the CBCT volume extended to cover the CT field of view. The second control point position matches the centre of the first voxel of the image.*

the contours from the pCT to the dCT, and the results saved in DICOM format. Both the dCT and warped structures were then imported back in the TPS for dose calculations.

A well-known issue with CBCT imaging is that the limited FoV often makes the images unusable for treatment planning due to missing patient information [9]. Methods proposed to handle this issue include acquiring two consecutive CBCTs [104], or directly using pCT slices to extend the CBCT [61]. The deformation outside the CBCT FoV was estimated by continuity, using the initial rigid alignment and the regularisation of the deformable registration. The CBCT volume was extended in the superior/inferior direction, to cover the whole CT FoV. The B-spline control point grid is placed on the CBCT by aligning the second CP with the centre of the first voxel. By "padding" the CBCT the B-spline control point placement is modified (Figure 3.1). The deformation outside the FoV was initialised using the rigid alignment. During the registration the transformation is optimised over the whole of the extended volume; as there is no image data to drive the registration outside the FoV it is purely driven by the constraint terms in these regions. This has the effect of causing a smooth transition between the image driven deformation inside the field of view and the rigid alignment outside the FoV. A good rigid alignment between the pCT and CBCT is then required, and it provides a good approximation mostly in the superior region, as the patient's head moves in a rigid fashion. The superior region is usually the most important due to presence of OARs such as brainstem and parotids. The brainstem only moves rigidly, but the parotids can shrink and migrate [43], and if not imaged the registration will likely represent the wrong deformation.

3.2.3 Evaluation of the suitability of deformable image registration for "dose of the day" calculations

Two independent tests were performed to assess the appropriateness of the proposed registration methodology for "dose of the day" calculations in a dCT. A geometric eval-

uation was performed to assess the ability of the proposed DIR method to map identical structures between the CT and CBCT datasets. Features delineated in the pCT were deformed using the output DVF and compared with the same features manually drawn on the CBCT. The results obtained were compared with those of using a rigid-only registration of the vertebrae. A dosimetric evaluation was performed to evaluate the impact of the registrations errors in the application proposed, to identify the limitations of the out-of-field approximation proposed, and assess how the method compares with other approaches. Dose distributions for the same IMRT plan were calculated on the dCT and rCT, and compared. For the second test the DIR results were compared not only with those from a rigid-only registration of the vertebrae, which approximates UCLH's current alignment protocol, but also with dose calculations directly on an calibrated and extended deformed CBCT (dCBCT) [97, 98].

3.2.3.1 Geometric evaluation

The purpose of this test was to assess the DIR ability to align the same anatomical features in CT and CBCT images. For each patient a set of easily identifiable features (Figure 2.10) was drawn by the same clinical expert on both the pCT and the CBCT, as described in Section 2.5.1.3. Typical HN OARs, such as parotids and brainstem, were not considered in this evaluation because they cannot be unequivocally seen in a CBCT scan. The CBCT used for each patient was the last acquired before replan referral.

The quantities calculated to assess the quality of the transformation for each structure were also previously described (section 2.5.1.3): DSC, OI, CoM and DT. For DT several statistics were used: the fraction of the DT distribution that is larger than 2 mm ($|DT|_{2mm}$), the signed and unsigned mean (DT_{mean} and $|DT|_{mean}$) and standard deviation (DT_{std} and $|DT|_{std}$), and the 95th percentile and maximum values of the unsigned DT ($|DT|_{95\%}$ and $|DT|_{max}$). While the signed DT provides a measure of bias (i.e., the registration being more likely to understimate or overestime contours), the unsigned DT provides an absolute measure of the variations. In addition, two other measures were also included:

- False positives (FP) as the fraction of deformed pixels (B) that are not part of the manual volume (A).

$$FP = \frac{|B \setminus A|}{|A|} \tag{3.1}$$

- False negatives (FN) as the fraction of manual pixels (A) that are not part of the deformed volume (B).

$$FN = \frac{|A \setminus B|}{|B|} \tag{3.2}$$

Using FP and FN as well as DSC provides additional insight into the cause of geometric errors, and will indicate if one structure is consistently larger/smaller than the other (Figure 3.2).

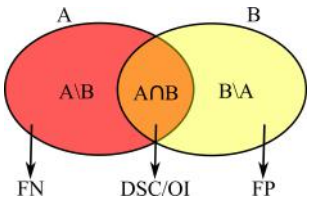

Figure 3.2: *False negatives (FN) and false positives (FP) versus dice similarity coefficient (DSC) and overlap index (OI).*

3.2.3.2 Dose comparison

In this second test the aim was to show that the pCT can be deformed into an image which is functionally equivalent to a rCT as far as dose calculation is concerned. To test this hypothesis, five different dose distributions were computed for each patient:

1. recalculated dose in the rCT (D_{rCT}), considered as gold-standard;
2. recalculated dose in the dCT, (D_{dCT}), the method here proposed;
3. recalculated dose in a rigidly aligned pCT (D_{pCT}), the current clinical approach;
4. recalculated dose in a not calibrated CBCT with superior/inferior extension ($D_{CBCT(nc)}$);
5. recalculated dose in a calibrated CBCT with superior/inferior extension (D_{CBCT});

In an ideal situation the rCT and CBCT would have been acquired at the same time, or at least in the same day, so that the two modalities contained the same (or comparable) geometric information. However using retrospective data such effort is not possible since there is no clinical reason to acquire a CT and a CBCT on the same day. The CBCT used for each patient was the one acquired on the first fraction after replanning. Since the rCT and following CBCT were not acquired simultaneously, but 5-7 days apart, noticeable changes to the patients' positioning and anatomy occurred between the scans. To minimise the errors in dose estimation due to discrepancies between the rCT and the CBCT, the pCT was actually registered to a dCBCT, obtained by deforming the regular CBCT to match the rCT. This dCBCT was closer to the ideal dataset discussed above. The rCT could have been deformed to match the CBCT instead but was not due to three reasons: (i) since the aim was to reproduce the dose calculated on a rCT it was preferable to not modify the rCT in any way, (ii) the plan isocenter was intrinsically defined in the rCT (further explanation on this point below), and (iii) since CBCT inherently has lower imaging quality possible errors in the dCBCT will be less noticeable than similar errors in deforming the rCT. All the quantitative results reported in the following sections refer to the use of dCBCT, and therefore the nomenclature dCBCT or CBCT is interchangeable.

To calculate D_{CBCT} the dCBCT values were replaced, pixel by pixel, with CT values and the image was extended in the superior and inferior direction using the corresponding rigidly aligned pCT slices. This approach was chosen in favour of defining REDs since the

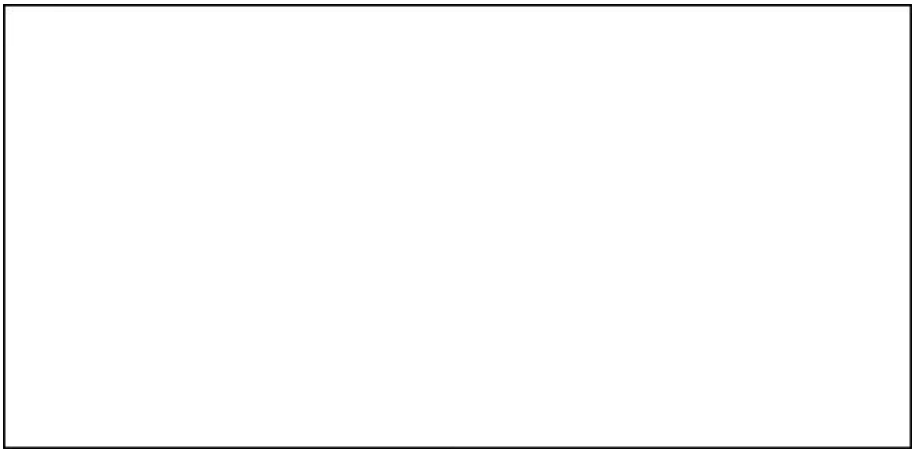

Figure 3.3: *Capthan 504 (a) position and composition of the inserts (adapted from [106]) and (b) image acquired with a CT scanner. Courtesy of James Annkah.*

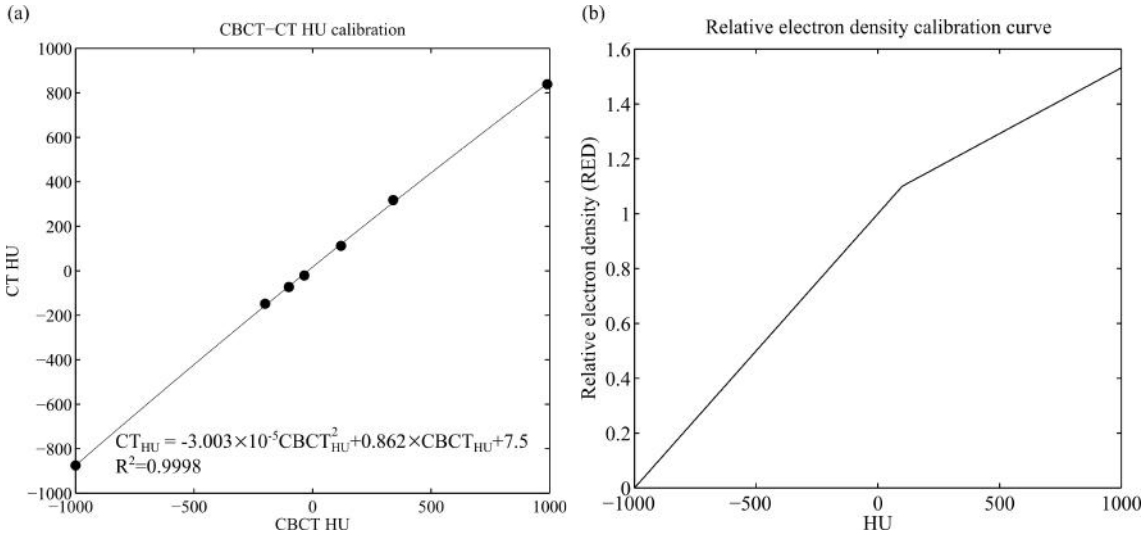

Figure 3.4: *(a) Relationship between CBCT and CT hounsfield units (HU) using the Catphan 504. The conversion curve between CT and CBCT numbers was approximated by a quadratic polynomial. (b) Relative electron density calibration curve for the CT scanner used in this study (CIRS Phantom).*

TPS does not allow to define multiple curves for the same dose calculation (i.e., one curve for the CBCT region, and other for the rigidly-aligned pCT). The relationship between CT and CBCT values was obtained using the Catphan-504 (The Phantom Laboratory, Greenwich, NY, USA) calibration phantom. Imaging and measuring of the phantom were performed by James Annkah (Department of Medical Physics & Biomedical Engineering, UCL) [105]. The average HU for each of its constituting materials was calculated for each of the imaging modalities (Figure 3.3a). The conversion curve between CT and CBCT numbers was approximated by a quadratic polynomial (Figure 3.4). Figure 3.5 shows the data used and registrations performed for each patient. $D_{CBCT(nc)}$ was calculated to verify if the calibration used did improve the dose calculations directly on the CBCT.

Doses were calculated for an IMRT plan using Eclipse (Varian Medical Systems, Palo Alto, CA, USA) External Beam TPS AAA with the highest available resolution (1 mm).

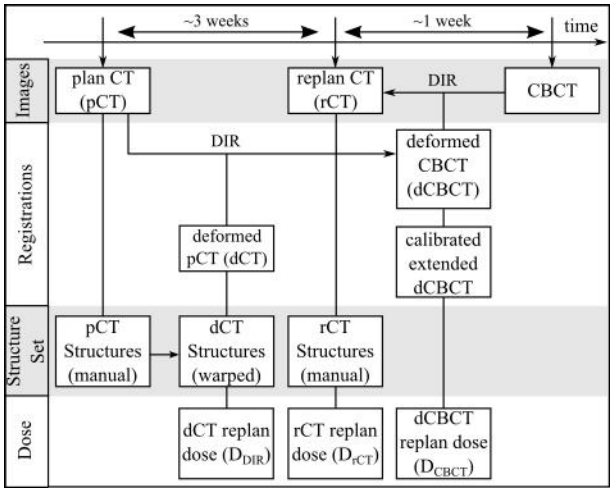

Figure 3.5: *Diagram of the data and registrations used in the dosimetric evaluation. The structures included in the dosimetric evaluation were the brainstem, spinal canal and parotid glands.*

For each patient the same IMRT plan, including beam arrangement, monitor units and fluence maps, were employed. The choice of IMRT plan should be clinically relevant, so a dose distribution the patient was treated with was chosen. The plan chosen to perform the dose calculations was the replan, which was built and optimised in the rCT. This minimises the issues with the isocenter definition both in the gold-standard, DIR-based dose calculation and dCBCT. The uncertainties with isocenter positioning are only an issue when calculating D_{pCT}. The isocenter uncertainty positioning in this case was dealt with by using the UCLH's protocol for isocenter alignment (i.e., matching to the spinal canal). The RED curve used in the dose calculations was the same used clinically at UCLH, which is yearly monitored using the CIRS phantom (Computerized Imaging Reference Systems Inc, Norfolk, VA, USA). The RED curve was fitted using a two-piece linear fit (Figure 3.4b).

The dose distributions were compared considering DD, gamma analysis, described in section 2.5.1.3, and similarity of the 95% isodose volume (representing target coverage).

The DD distributions were compared in terms of test pass-percentage (i.e., percentage of pixels whose DD was inferior to 2%pD ($DD_{2\%\text{-}pp}$), average absolute DD mean and root mean square (RMS) (DD_{mean} and DD_{RMS}), and the 99th percentile of the DD distribution ($DD_{99\%}$).

3.2.3.3 Propagation of structures and "dose of the day"

The dose analysis was extended to examine the impact within different OARs. In clinical settings, DVHs are routinely used to assess if the plan is appropriate for the patient, by displaying in a concise and comprehensive way the information of dose delivered both to targets and OARs. DVHs were computed using both manually drawn and warped structures of interested. The structures were delineated by the radiographers as part of clinical

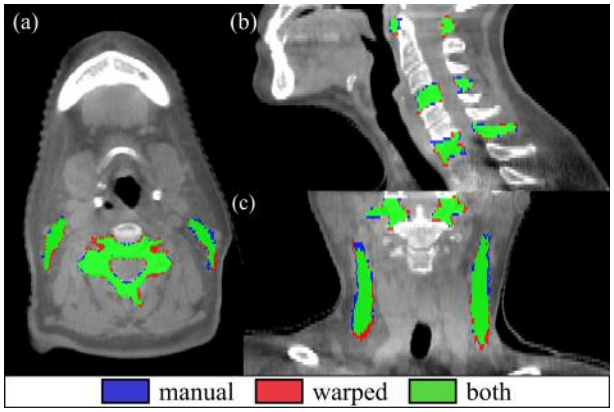

Figure 3.6: *Geometric matching of manual and warped features overlaid on the CBCT image. Blue corresponds to manual-only, red to warped-only voxels, and green to the region of agreement. The features shown are (from left to right): (a) right sternocleidomastoid muscle (RSCM), C4 and left sternocleidomastoid muscle (LSCM), (b) vertebrae C1, C4 and C7; and (c) RSCM, C1 and LSCM.*

practice. Complementary to DVHs analysis, overlap between OARs manually drawn in the rCT and warped from the pCT was also assessed and the differences predicting the mean (ΔD_{mean}) and maximum doses (ΔD_{max}) to OARs were computed.

3.3 Results

3.3.1 Geometric evaluation

Figure 3.6 shows a representative example of the matching of manual and warped features. The visual matching of the features is satisfactory after registration, particularly for complexly shaped features such as the vertebrae.

In Table 3.2 the mean DSC, OI, FN, FP, DT and CoM obtained for different types of features is presented. Figure 3.7 shows the complete information of the distribution of DT values, for different feature and registration types. DIR aligns the features well and considerably better than rigid registration, and the results obtained are more consistent between different structures types. The results obtained are also poorer in soft tissue region than in bone anatomy. Inherent lower soft tissue contrast in the CBCT degrades both the registration accuracy and the quality of manual segmentations. A major improvement in FN and FP can be found when using DIR, particularly for soft tissues. The tail of the DT distribution, and consequently maximum DT values, are thought to be more related with local poor manual segmentation than to registration errors. The positive values of DT_{mean} are indicative that the registration is slightly biased toward overestimating the size of the structures. Since in the HN cohort in general the patients lose weight, it means the registration is more likely to not be able to fully capture the contraction of the tissues.

Table 3.2: *Mean values ± standard deviation of dice similarity coefficient (DSC), overlap index (OI), false positives (FP), false negatives (FN), distance transform (DT) and centroid position error (CoM) obtained using deformable (DIR) and rigid-only (RIG) registrations. The results were presented for all structures/patients, and also grouped by different structure type: external contours, bony anatomy (vertebrae C1, C4 and C7), and soft tissues (left and right sternocleidomastoid muscles).*

	Externalcontours		Bony anatomy		Soft tissues		Overall	
	RIG	DIR	RIG	DIR	RIG	DIR	RIG	DIR
DSC	0.945 ±0.017	0.986 ±0.001	0.72 ±0.12	0.85 ±0.03	0.64 ±0.14	0.79 ±0.06	0.73 ±0.16	0.85 ±0.08
OI	0.983 ±0.007	0.988 ±0.003	0.74 ±0.12	0.89 ±0.04	0.68 ±0.12	0.80 ±0.06	0.76 ±0.15	0.88 ±0.08
FP	0.10 ±0.04	0.016 ±0.003	0.31 ±0.15	0.21 ±0.05	0.5 ±0.2	0.23 ±0.08	0.3 ±0.2	0.19 ±0.09
FN	0.015 ±0.006	0.012 ±0.003	0.25 ±0.11	0.11 ±0.04	0.28 ±0.08	0.19 ±0.06	0.22 ±0.13	0.12 ±0.08
DT_{mean} (mm)	1.5 ±0.8	0.13 ±0.09	0.7 ±0.5	0.43 ±0.15	1.2 ±0.9	0.52 ±0.18	1.0 ±0.8	0.4 ±0.3
DT_{std} (mm)	2.8 ±1.1	1.3 ±0.4	1.2 ±0.4	0.9 ±0.17	1.8 ±0.8	1.1 ±0.4	1.7 ±0.9	1.0 ±0.4
$\|DT\|_{mean}$ (mm)	1.9 ±0.8	0.4 ±0.1	1.4 ±0.6	0.8 ±0.1	1.9 ±1.0	1.0 ±0.2	1.6 ±0.8	0.8 ±0.3
$\|DT\|_{std}$ (mm)	2.8 ±1.1	1.3 ±0.4	1.2 ±0.4	0.9 ±0.2	1.6 ±0.6	1.1 ±0.4	1.6 ±0.8	1.0 ±0.4
$\|DT\|_{95\%}$ (mm)	8 ±3	2.1 ±0.3	3.5 ±1.4	2.4 ±0.6	5 ±2	2.9 ±1.1	5 ±3	2.5 ±0.9
$\|DT\|_{max}$ (mm)	22 ±9	19 ±8	7 ±2	5.5 ±1.6	10 ±4	9 ±4	10 ±7	9 ±6
$\|DT\|_{2mm}$ (%)	31 ±9	3.6 ±1.5	20 ±18	5 ±4	29 ±22	9 ±5	25 ±19	6 ±4
CoM (mm)	4.4 ±1.4	0.8 ±0.3	2.5 ±1.6	0.8 ±0.4	4 ±3	2.1 ±0.6	3.2 ±2.2	1.2 ±0.8

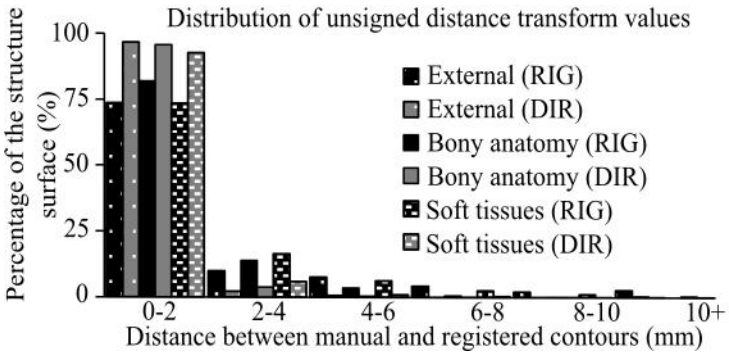

Figure 3.7: *Distribution of distance transform (DT) values for deformable (DIR, in grey) and rigid only (RIG, in black) registrations. The results are grouped by different structure type: external contours, soft tissues (left and right sternocleidomastoid muscles) and bony anatomy (vertebrae C1, C4 and C7).*

3.3.2 Dose comparison

Overall D_{dCT} matches D_{rCT} well (Figure 3.8). The dose similarity results were analysed based on different regions of the patient (Table 3.3), providing evidence of the dose behaviour outside the CBCT FoV. D_{dCT} results were better than the D_{pCT} and D_{CBCT} results in all regions, although the major benefits of using D_{dCT} were in the region imaged in the CBCT where most anatomical changes occur and higher dose is delivered. Doses similarity statistics on a calibrated CBCT were superior to those on a non-calibrated CBCT, showing that machine-specific calibration improved the usability of CBCT for direct dose calculations.

Target coverage similarity was then assessed in terms of the 95%-isodose volumes obtained for D_{dCT}, D_{pCT} and D_{CBCT} in comparison with D_{rCT} (Table 3.4). There is a good agreement with the rCT in terms of similarity of the isodose curves obtained when using the dCT.

The DD inside the OARs are most relevant clinically (Figure 3.9). In the brainstem all methods behave similarly since all are based, directly or indirectly, in information from the rigidly aligned pCT. D_{dCT} DD values were in general clinically insignificant inside OARs, and the results obtained are superior to those from other dose estimation approaches. The poor image quality of the CBCT in the inferior direction (i.e, larger imaging volumes such as the shoulders) is responsible for the inferiority of D_{CBCT} in comparison with D_{pCT}. This affected both the dose estimation in the high dose region (Table 3.3 and Figure 3.8) and in the spinal canal (Figure 3.9b). The inconsistency in HU is less problematic in the remaining OARs.

3.3.3 Propagation of structures and "dose of the day"

Figure 3.10 contains DVHs calculated using (i) D_{rCT} and manually drawn structures on the rCT, (ii) D_{dCT} and the same manual structures, and (iii) D_{dCT} and structures warped

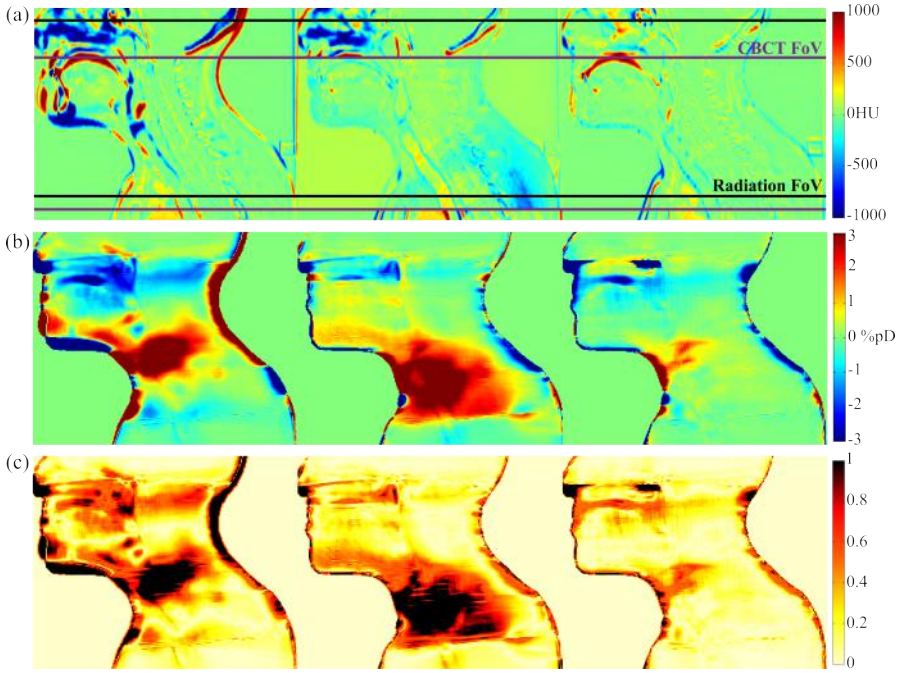

Figure 3.8: *Rigid (left), extended CBCT (middle) and deformable registrations (right) (a) intensity difference image with the replan CT (b) dose difference with D_{rCT} as percentage of the prescribed dose (%pD) and (c) gamma analysis. Between 60-90% of the treatment field-of-view was imaged in the CBCT in the available datasets. Treatment (black line) and CBCT (purple line) fields-of-view are indicated in (a). Deformable results are clearly better than rigid and extended CBCT. Most striking registration errors, and therefore dose estimation errors, occurred in the skin and airways. The inconsistency in HU is visible for the CBCT results, particularly in the shoulder region. This degrades the accuracy of the dose estimation in a fraction of the high dose region and spinal canal.*

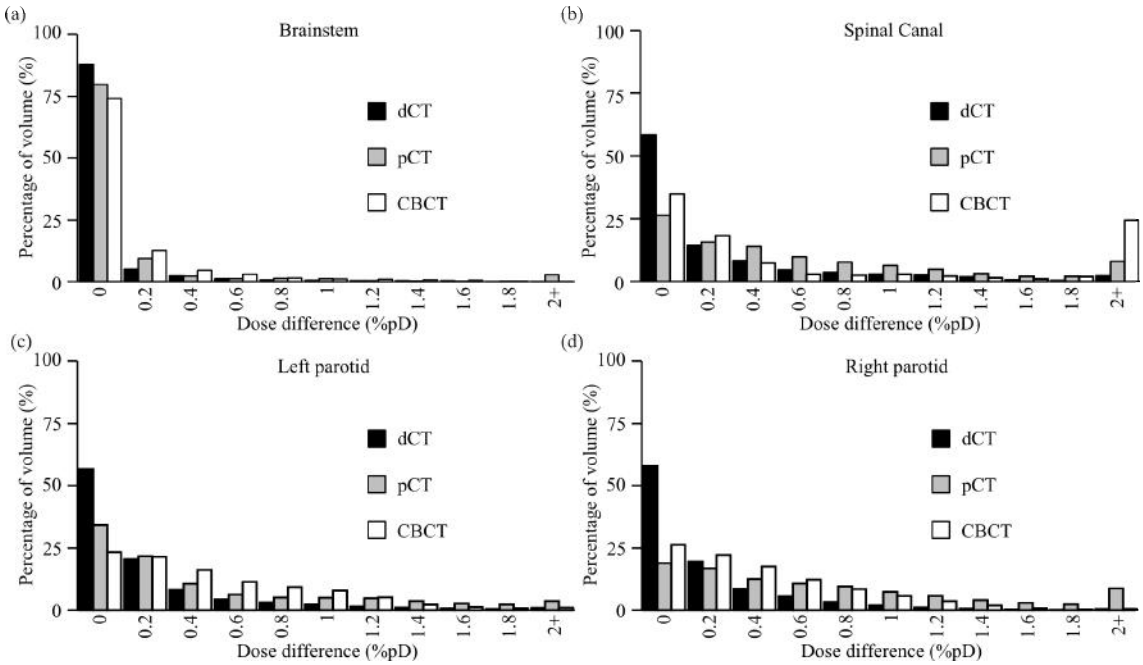

Figure 3.9: *Dose difference with D_{rCT}: distribution of values (as percentage of the prescribed dose, %pD), for different regions of interest and dose calculations (D_{dCT}, D_{pCT} and D_{CBCT}).*

Table 3.3: *Similarity between dose distributions [deformable (D_{dCT}), rigid-only (D_{pCT}), extended CBCT [(calibrated (D_{CBCT} and non-calibrated ($D_{CBCT(nc)}$)] and gold-standard (replan CT) within different regions of interest: mean ± standard deviation for DD pass-percentage with a tolerance of 2%pD ($DD_{2\%-pp}$), mean (DD_{mean}), root-mean square value (DD_{RMS}) and 99th percentile ($DD_{99\%}$) and gamma test pass-percentage (2%/2mm criterion).*

		Dose Difference (DD) test				Gamma test
Region of interest	Method	$DD_{2\%-pp}$ (%)	DD_{mean} (%pD)	DD_{RMS} (%pD)	$DD_{99\%}$ (%pD)	Pass-percentage (%)
Treatment FoV	D_{dCT}	87±6	1.6±1.0	5±2	25±12	94±5
	D_{pCT}	76±4	3.9±0.8	10.9±1.4	55±7	85±3
	$D_{CBCT(nc)}$	75.3±1.7	2.2±0.5	5.5±1.9	26±11	88.7±1.9
	D_{CBCT}	78.6±1.0	2.0±0.5	5±2	24±12	91±3
CBCT FoV	D_{dCT}	90.0±0.9	1.2±0.2	4.5±1.0	22±6	97.1±1.1
	D_{pCT}	74±3	4.4±0.8	12.1±1.6	59±8	84±2
	$D_{CBCT(nc)}$	75±10	1.9±0.6	4.4±0.8	20±4	90±5
	D_{CBCT}	80±8	1.6±0.5	3.7±0.5	17±2	93±4
Non-imaged treatment FoV	D_{dCT}	86±11	1.8±1.5	5±3	26±17	92±8
	D_{pCT}	84±9	2.1±1.3	6±3	29±17	90±7
	$D_{CBCT(nc)}$	84±9	2.0±1.3	5±3	28±17	90±7
	D_{CBCT}	84±9	2.0±1.3	5±3	28±17	91±7
Treated Volume	D_{dCT}	93±8	0.7±0.2	1.2±0.5	4±2	97±4
	D_{pCT}	68±7	2.1±0.6	5±4	29±37	80±6
	$D_{CBCT(nc)}$	81±5	1.5±0.3	2.1±0.3	6.2±0.6	88±5
	D_{CBCT}	84±5	1.2±0.2	1.8±0.2	5.4±0.8	90±4

Table 3.4: *Similarity between the isodose volumes [deformable (dCT), rigid-only (pCT) and calibrated CBCT], and the gold-standard (replan CT): mean ± standard deviation (and range) of dice similarity coefficient (DSC), overlap index (OI), fraction of the distance transform distribution larger than 2 mm (DT_{2mm}) and centroid position error (CoM).*

Method	DSC	OI	DT_{2mm} (%)	CoM (mm)
dCT	0.962±0.015 (0.937-0.978)	0.974±0.007 (0.963-0.982)	5±6 (0.2-14.4)	1.1±1.0 (0.2-2.9)
pCT	0.929±0.016 (0.908-0.950)	0.930±0.028 (0.884-0.958)	20±5 (12.1-26.8)	2.7±0.4 (2.2-3.2)
CBCT	0.957±0.011 (0.940-0.971)	0.979±0.007 (0.968-0.986)	14±4 (10.6-20.7)	1.6±0.2 (1.3-1.9)

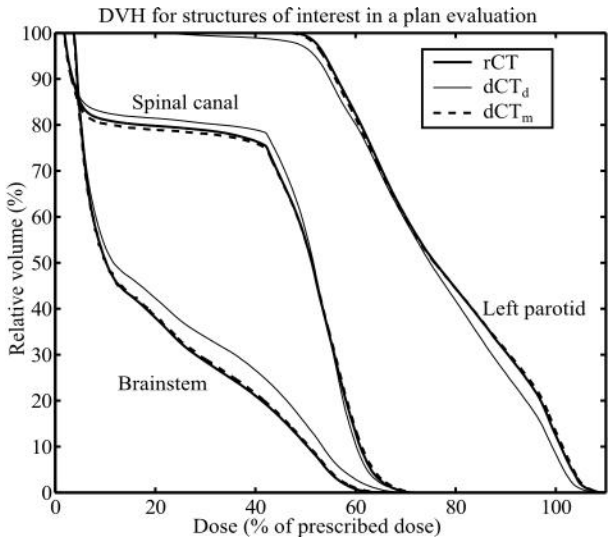

Figure 3.10: *Example of a DVH for different OARs using D_{rCT} and manually drawn structure (rCT), D_{dCT} and structures deformed from the pCT (dCT_d), and D_{dCT} and manually drawn structures (dCT_m). Right parotid omitted for figure clarity.*

from the pCT. Measures of overlap obtained for OARs show that although the volumes are similar the differences can be non-negligible in the DVHs. The mean DSC obtained was 0.81±0.04, 0.82±0.06, 0.76±0.05, 0.770±0.010, for the brainstem, spinal canal, left parotid and right parotid respectively. From Figure 3.10 it can be seen that differences in DVHs were mainly due to differences in the OARs definition (manual vs warped), and not due to differences in dose estimation (D_{rCT} vs D_{dCT}). The same trend was found for the errors in estimating the maximum and mean doses to an OAR (ΔD_{max} and ΔD_{mean}). Using dCT combined with manually drawn structures on the rCT the mean value obtained for ΔD_{mean} was 0.1±0.1% of the pD (range: 0.0-0.3%pD), while for ΔD_{max} was 0.3±0.2%pD (range: 0.0-0.6%pD). However, using the warped structures these errors increased to 2.4±2.1%pD (range: 0.3-7.8%pD) and 1.5±1.6%pD (range: 0.1-6.2%pD), respectively.

3.4 Discussion

Regarding the geometric evaluation, all metrics showed an improvement when comparing deformable to rigid registration, up to a relative improvement of 80%. The values found for DSC were comparable to the ones obtained by Castadot et al. using CT-to-CT DIR [87]. DT metrics and CoM values were comparable to image resolution. Combining the DSC, OI, FN, FP, DT and CoM the ability of the DIR method to map identical structures between the CT and CBCT datasets could be evaluated. However, each of these quantities can be misleading on its own. For example, DSC is inherently bigger for larger structures, such as the external contours, and in this case DT is a better measure of similarity between features.

One of the main limitations of the geometric evaluation was not including the localisa-

tion of anatomical landmarks. Anatomical landmarks were not used in this study due to the difficulty detected by the clinical expert in consistently identifying points in a CBCT image. The uncertainty in landmark locations could produce misleading accuracy results, so the option chosen was to only use structures in the geometric evaluation. Furthermore, for the current goal of calculating the "dose of the day" accurate point-to-point mapping is not required.

The method presented allowed for accurate dose calculations, comparable to doses recalculated on a rCT and superior to both the current clinical approach at UCLH and dose calculations on the CBCT (with extended FoV). CBCT images include larger amounts of scattering than CT, resulting in larger variation in HU values that limit the HU calibration and reliability [49]. The CBCT calibrations are usually done using a small phantom which provides consistent results in such a small FoV. However, for larger volumes the calibration is no longer consistent which has a considerable impact in dose calculations. Figure 3.8 showed the effect in dose estimation of such inconsistency in HU in the neck and shoulders region. Imaging larger volumes result in increased scatter and reduced transmission. The increase in scatter introduces non-uniformities and additional quantum noise to the reconstructed image [107]. This indicates the need for more specific and appropriate calibration phantoms for CBCT, which should cover the size of the treatment region. The choice of phantom is crucial as others showed different phantoms result in very different RED calibration curves, and particularly the Catphan may not be the most appropriate due to issues with its bone-equivalent material [99]. While the results found are indicative of how reliable CBCT currently is for direct dose calculations, it was not in the scope of this work to optimise treatment planning on CBCT images. The imaging protocol was not optimised for that purpose, and so the calculations were clearly suboptimal. It was verified that calibrating the CBCT improved all the results when compared with non-calibrated CBCTs. This improvement was considerable but not enough, informing that artefacts and scatter have a large impact that a single calibration curve cannot recover. Calculating dose distributions directly on CBCT images is still an active area of research and the fact that specialised calibrations and optimisations are required is a current disadvantage of such methods. The deteriorated image quality of CBCT leads to serious concerns about its reliability for direct dose calculations. CT is still far superior to CBCT for treatment planning, and DIR is a good interim solution for ART until CBCT data is directly usable.

The results presented were promising even to obtain dosimetric information outside the CBCT FoV. The interpolation of the information outside the FoV allowed performing dose calculations even when the CBCT FoV was smaller than the treatment FoV. The method proposed to estimate the transformation, and hence the anatomy, outside of the CBCT FoV will however not always give trustworthy results. If there is significant deformation outside the FoV then it will not be able to recover this. Such deformations occurred for one of the patients included in this study where the CBCT FoV did not

extend far enough in the inferior direction. Further research will be required to study the validity of this approximation. For future clinical applications, the imaged region must be selected properly to minimise possible out-of-field errors on critical regions of the individual patient. For example, if the major concern is the dose given to the brainstem, then the brainstem should be imaged. If target coverage is more important, the high dose region should be properly captured. Informing the imaging procedures will be even more important for bigger patients and/or tumours.

Figure 3.10 shows how structure delineation has an important impact in plan evaluation. Structure contours deviations explain the differences in DVHs and mean/maximum doses to OARs. The overlap found between OARs is similar to values reported by Tsuji et al. for CT-to-CT DIR [51]. Visually most discrepancies in spinal canal and brainstem were actually due to differences in defining the extent of these organs in the superior direction in the different scans. Generating appropriate structure delineations for ART is an important and on-going area of research [10, 51, 54], which is beyond the scope of this work. In future applications of the tool presented and validated here, deformed structures will be used as a starting point to speed up the evaluation process, but will likely require manual verification and editing to be used clinically. Extra care must be taken if modifying targets during ART. Even though authors suggest the use of DIR to monitor tumour shrinkage [55], in CT-to-CBCT DIR warping target volumes may not be appropriate as they are usually not visible in CBCT imaging. Also, even if the GTV shrinkage is visible, there is no evidence that microscopic proliferation has shrunk in the same proportion. Guidelines for target propagation during ART need to be developed. Including routine functional imaging, such as MRI, in the ART workflow may provide not only a solution for target propagation, but also early evidence of the patient response to the treatment [108, 109]. The topic of multimodal and functional imaging will be the focus in chapter 7.

Further work on this topic should focus on measuring the accuracy of the deformation maps and further improvements of the registrations. One of the sources of errors in DIR is the inherent deformation of bony elements, which physically can only move rigidly. Rigidity penalty terms, that constrain the registration to be rigid in regions of interest, are desired in a realistic deformable registration algorithm to increase the accuracy of the tissue mapping [110]. Other similarity measures can also be investigated to improve the robustness of the registrations [29, 111]. In chapters 4 and 5 other algorithms and similarity measures are investigated for dose warping and summation applications, as well as proton therapy.

The patients included in this study had considerable anatomical changes during the course of their treatment. Therefore the tests applied to the registration algorithm were quite severe as those patients were selected from the group of identified sensitive cases treated at UCLH. The registrations were particularly challenging for the dose comparisons since not only anatomical changes but also different positioning systems had to be

reproduced. For this reason it is expected that routine clinical cases will be less demanding.

The initial clinical application of the method presented in this chapter will be weekly offline calculations of the "dose of the day" to help inform the decision of whether the current plan is still acceptable. At this point if a plan is found to be unacceptable the current replan pathway will be followed (i.e., acquire a rCT and replan). Replanned patients will be used on further validation to support the effectiveness and efficiency of the proposed method. With more patients it may be possible to understand the relationship between DIR and dose errors, which could be used to establish quick and easy methods for detecting regions where DIR errors are significant from a dose calculation point-of-view. On a longer term the aim is to use dCT directly in the replan procedure itself and to remove the need for acquiring a new CT when a new immobilisation is not necessary. The final aim is to enable the implementation of a controlled "dose-driven" ART approach that can be built into the patient pathway: to perform routine online modifications to the treatment plan based on the dose that has already been delivered.

3.5 Conclusions

This work presented a proof-of-principle of the application of an in-house developed deformable registration for ART purposes. CT-to-CBCT DIR was developed, optimised and evaluated, and it was demonstrated that using a pCT scan deformed to match a CBCT scan resulted in similar dose calculations to those performed on a new CT scan. The DDs were clinically acceptable, and DIR and CBCT-based dose calculations provided an estimation closer to the gold-standard than calculations in a rigidly aligned pCT and extended CBCT. The results obtained support the use of non-rigid registration and provide further evidence in the challenging objective of validating deformable registration for routine clinical use. The registration methodology and validation protocol were implemented in a friendly and semi-automated fashion using MATLAB (MathWorks, Natick, MA, USA), and was made available in a clinical workstation at the Radiotherapy Department (UCLH) to be used in evaluating the need to replan newer patients and further clinical validation.

Chapter 4

Dose warping and summation applications

> Measure what can be measured, and make measurable what cannot be measured.
>
> Galileo Galilei

In the previous chapter the use of a dCT as a platform for ART, particularly dose recalculation, in the context of IMRT treatments was investigated. This chapter focuses on dose warping and summation applications. The evaluation of more recent and sophisticated registration algorithms available in NiftyReg was a major point of the current chapter.

The work in this chapter resulted in the following outputs:

- C. Veiga, A. Lourenço, S. Moinuddin, M. van Herk, M. Modat, S. Ourselin, D. D'Souza, G. Royle, and J. R. McClelland, "Toward adaptive radiotherapy for head and neck patients: uncertainties in dose warping due to the choice of deformable registration algorithm," Med. Phys. 42(2) 760-769 (2015).
- A. M. Lourenço, C. Veiga, G. Royle, and J. McClelland, "Dose remapping and summation for head and neck adaptive radiotherapy applications", NPL PPRIG Proton Therapy Physics Workshop (London, United Kingdom, 2014).

4.1 Rationale

The concept of dose warping stands for the clinical application of using DIR to map the dose delivered by a treatment between different time points. Mapping the dose from different treatments at different time points allows for dose summation, i.e., knowing the

total dose deposited over a period of time when the anatomy is varying. Dose warping is therefore sensitive to the underlying deformation, and algorithms that ensure symmetry, inverse-consistency and diffeomorphism (concepts introduced in section 2.2.4) may be preferable over more traditional implementations.

Dose warping and summation is an important topic in ART research, since knowledge of the total dose delivered at each time point is fundamental in the ART decision-making and replanning process. Validation of DIR is not only a key aspect, but also one of the most challenging; therefore some authors have been evaluating the accuracy of DIR and dose warping using manually annotated points and structures [10, 54, 112, 113], physical plausibility of the transformations [114, 115] and by developing deformable and virtual phantoms with known deformations [116–119]. Others focused on estimating the accuracy requirements for dose warping [120, 121] or estimating its precision [39, 122–124]. Monte Carlo methods have also been proposed to recalculate doses using a deformed grid (with deformed and irregular voxels), which still uses DIR but minimises the errors associated with dose interpolation [125–127]. Finally, some groups have been using dose warping to evaluate the benefits of replanning [55, 56, 128]. Most of the work done on these topics uses CT-to-CT registration on different anatomical sites, and/or registration algorithms that do not ensure symmetry and inverse-consistency.

In this chapter an ART framework for HN patients using the CT and CBCT imaging was investigated. The HU information is mapped to the daily geometry for "dose of the day" calculations, and the dose is remapped to the planning geometry for dose summation. Therefore estimates of both forward and inverse transformations are required. Other groups suggest using CT and CBCT for dose remapping without requiring both transformations by calculating the "dose of the day" directly in the CBCT [55, 113]. However, and as seen in chapter 3, CBCT imaging is still unreliable for direct dose calculations and until the image quality of CBCT is improved a dCT is a good interim solution. Four approaches to obtain those transformations were tested and evaluated, using three different DIR algorithms implemented in NiftyReg, all of which use a B-spline parametrisation of the transformation with different theoretically desirable properties. The aims of this work were to compare different DIR algorithms available in NiftyReg in terms of their geometrical accuracy, physical properties and computation time, and to evaluate the uncertainties inherent to using different algorithms for dose warping. Since algorithms from the same software package (i.e., similar implementation) are compared, the differences found are due to the physical properties of the underlying algorithm and not due to other differences in implementation.

The preliminary work on this topic must be acknowledged to Ana Mónica Lourenço (Department of Medical Physics & Biomedical Engineering, UCL), whom as part of her MSc thesis investigated different methods to numerically calculate the inverse of a DVF, how to evaluate the inverse-consistency and how to use NiftyReg for dose warping [129]. I was closely involved in the supervision of her work.

4.2 Methods and Materials

4.2.1 Patient data acquisition

A total of five head and neck patients were used in this study. This was the same cohort described in sections 2.5.1.1 and 3.2.1.

4.2.2 Image registration settings

The different algorithms available in NiftyReg were used: standard asymmetric, inverse-consistent symmetric and velocity fields implementations. BE was used as a regularisation term for all of the registrations [21]. In addition JL [103] was used for the unidirectional registrations and the IC [36] was used for the inverse-consistent registrations. The penalty term weights were optimised for each algorithm independently. The remaining parameters and approaches used (such as out-of-field approximation) to register the images are in agreement with the methods described in chapters 2 and 3.

The DIR algorithm used in this study was an updated version of the code used in chapter 2. NiftyReg was constantly updated throughout the duration of this thesis, and the code suffered alterations that were reflected in quality of the registrations obtained. The newer code was tested and evaluated with the same criteria described in the previous chapters. The conclusions derived previously are still true at this point, and the newer version was found to have improved performance in terms of propagation of structures and accuracy of the dose calculations.

4.2.3 Dose warping and summation in an adaptive radiotherapy workflow

The method proposed to compute, map and accumulate dose distributions while accounting for anatomic variations requires the pCT and CBCTs images acquired throughout the treatment. The process consists of each of the following steps (Figure 4.1):

(i) Mapping the HU information from the pCT to the CBCT geometry using DIR. The process is repeated for each CBCT available for the patient.

(ii) Dose calculations are performed on each dCT ("dose of the day").

(ii) Mapping the "dose of the day" back to the space of the pCT.

(iv) Dose distributions are accumulated and displayed on the pCT space. When a CBCT is not available for every fraction of the treatment, the weighting used when summing the doses will depend on the fractionation scheme and scans available.

The accumulated dose can potentially be used clinically to feed the ART decision-making process. At each imaging time point the dose delivered is mapped to the same

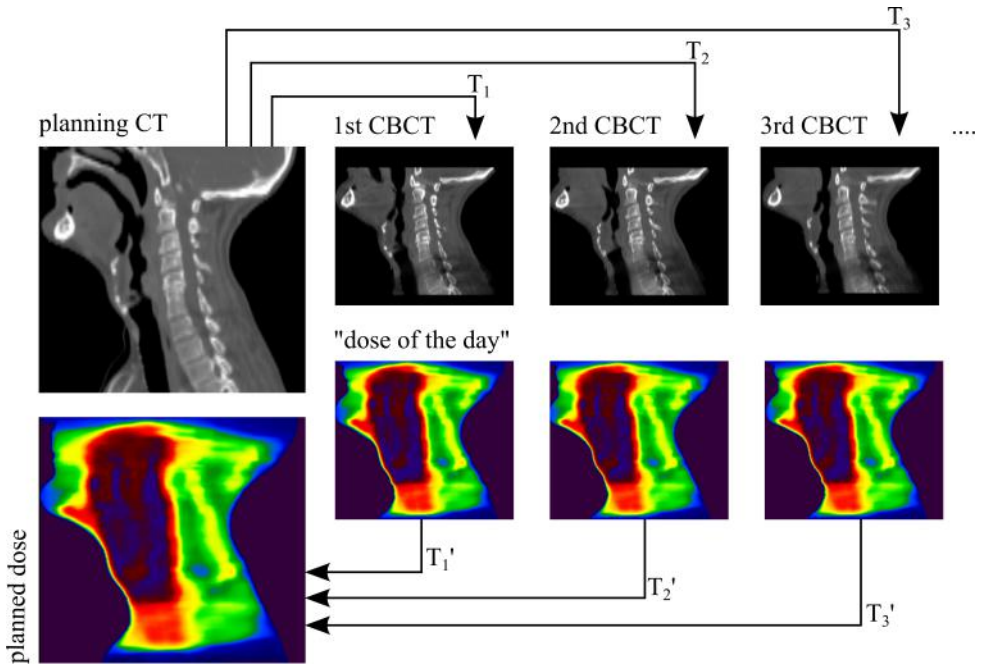

Figure 4.1: *The use of dose warping and summation in an adaptive radiotherapy workflow. The registration maps the Hounsfield units from the CT to the daily CBCT scans, and the deformed CT is used for "dose of the day" calculations. The dose delivered is mapped back to the planning stage, where it is accumulated and the need to replan assessed.*

reference space, chosen to be the pCT, where it is summed to the dose from previous time points. Choosing the pCT as reference allows to iteratively compare the planned objectives with the delivered values as the treatment progresses, such that the need to replan can be regularly assessed [55, 56]. The total dose delivered can also be warped to the daily geometry to assess the need to replan (using the CBCT as reference) and to feed the replanning process (using a rCT). In the first case additional registrations would be required between the fractions since the reference space would change at each fraction. The latter process should be performed with utmost care as uncertainties in registration will not only affect the decision to replan but also the newly planned treatment.

Four different approaches to dose warping available with the NiftyReg software were compared ("forward direction" stands for the CT-to-CBCT transformations and "backward direction" stands for the CBCT-to-CT transformations):

1. Standard asymmetric registration in both forward and backward direction ($DIR_{sas+sas}$);

2. Standard asymmetric registration in the forward direction followed by the numerical estimation of the inverse of this transformation ($DIR_{sas+inv}$);

3. Inverse-consistent symmetric registration which provides both the forward and backward transformations (DIR_{ics});

4. Symmetric registration parameterised by a stationary velocity field which inherently

Table 4.1: *DIR mapping approaches used in this study, and theoretical properties of the algorithms in terms of symmetry, inverse-consistency and diffeomorphisms.*

	Bidirectional	Symmetric	Diffeomorphic	Inverse-consistent
$DIR_{sas+sas}$	No	Yes	Yes (by constraint)	No
$DIR_{sas+inv}$	No	No	Yes (by constraint)	Yes
DIR_{ics}	Yes	Yes	Yes (by constraint)	Yes (by constraint)
DIR_{svf}	Yes	Yes	Yes (by implementation)	Yes

provides both the forward and backward transformations (DIR_{svf})

All the algorithms use a B-spline parametrisation of the transformation with different theoretically desirable properties: while $DIR_{sas+sas}$ is symmetric but not inverse-consistent, $DIR_{sas+inv}$ is inverse-consistent but not symmetric. DIR_{ics} and DIR_{svf} are both symmetric, but while the first encourages inverse-consistency using a penalty term, the second guarantees it by using a velocity field parametrisation [37]. Table 4.1 describes the theoretical properties of each of the parametrisations, in terms of directionality (i.e., if T^{-1} is generated explicitly), symmetry, inverse-consistency and diffeomorphisms.

4.2.4 Evaluation scheme

The different DIR algorithms were evaluated and compared in terms of geometrical matching, properties of the underlying deformations and computation time. All the statistical tests performed on this chapter were done using the MATLAB (MathWorks, Natick, MA, USA) statistics toolbox and the Wilcoxon signed ranked test (95% confidence).

4.2.4.1 Geometric matching

The purpose of this part of the study was to assess the ability of the different approaches to align the same anatomical features in CT and CBCT images, using the following quantitaties: DSC, OI, FN, FP, DT and CoM (section 2.5.1.3 and 3.2.3.1); and the structure set consisting of the external contours, vertebrae C1, C4 and C7, RSCM and LSCM (section 2.5.1.3). Since the registrations were evaluated in both direction, care was necessary when calculating the signed DT to interpreter correctly the results. Therefore, in the forward direction DT is negative if the deformed contour is inside the manual contour, otherwise it is positive (and vice-versa for in the backward direction).

4.2.4.2 Characteristics and similarity of the deformation fields

By optimising properly the parameters used in the different algorithms, it should be possible to obtain similar geometric matching. However, different algorithms can result

in very different DVHs, particularly inside anatomy that lacks internal features or regions of increased noise and reduced contrast in the CBCT scans. Therefore, characteristics of the underlying deformations were evaluated in this section.

The smoothness of the transformations was analysed using the harmonic energy (HE) and the properties of the determinant of the Jacobian of the transformation [det(Jac)]. The HE refers to the mean Frobenius norm of the displacement field, and is inversely proportional to the smoothness of the deformation [115]. det(Jac) indicates the level of expansion/contraction at each voxel, and negative values are indicative of noninvertible and unrealistic deformations. Additionally, the ICE was calculated to investigate if the transformations were true inverses. Considering that the forward and backward transformations are inverse-consistent if their composition is equal to the identity transform, the ICE was calculated as the voxelwise difference between the composed transformations and the identity transform. The code used to calculate the ICE was implemented Dr. Jamie McClelland (CMIC, UCL) and benchmarked by Ana Mónica Lourenço (Department of Medical Physics & Biomedical Engineering, UCL).

Additionaly, the similarity between mappings using different algorithms was assessed by measuring voxel-by-voxel the Euclidean distance between the DVFs (computed as a L^2-norm). This measure is indicative of the variability in mapping between different DIR algorithms.

4.2.4.3 Computation times

The time taken to complete the registrations is important for clinical translation, therefore the computation times for each approach were measured three times per dataset. Non-GPU implementations were done on an Intel® Core™ i7-2600S CPU (2.80GHz, 8GB RAM), and the GPU registrations used a NVIDIA Tesla C2070 GPU card (14 multiprocessors, 6GB dedicated memory).

4.2.4.4 Dose warping comparison

The purpose of this part of the study was to investigate the uncertainties in dose warping when using different DIR algorithms.

Dose calculations were performed on a dCT, and mapped back to the original pCTs using the results from the four different methods (in all cases cubic spline interpolation was used). Varian Eclipse External Beam Planning System AAA was used to calculate the dose distributions using the highest available resolution (1 mm). For each patient the same IMRT plan was applied, including beam arrangement, monitor units and fluence maps. The dose distributions were compared within different volumes of interest using DD, by calculating the percentage of pixels whose DD was inferior to 2% of the prescribed dose (pD) ($DD_{2\%\text{-pp}}$), mean value (DD_{mean}), RMS value (DD_{RMS}), and the 99^{th} percentile of

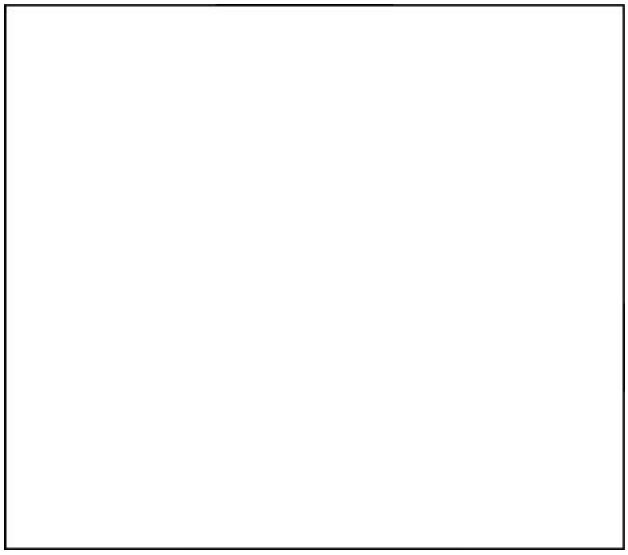

Figure 4.2: *Basic flow diagram of the distance to dose-difference (DTD) algorithm, with a distance to agreement (DTA) tolerance level of δ. Adapted from [120].*

the DD distribution ($DD_{99\%}$); the differences predicting the mean (ΔD_{mean}) and maximum doses (ΔD_{max}) to OARs and dose volume histograms (DVHs).

The distance to dose-difference (DTD) was also calculated for all the planned dose distributions (using an accuracy of 2%pD) [120]. DTD is a method to estimate the required spatial accuracy of a DVF for dose warping based on the distance that one has to travel within the dose map to find a DD above a tolerance value. A DTD algorithm uses a single dose distribution, and determines the minimum distance one must traverse in this dose distribution to observe a DD greater than the tolerance (δ). The DTD algorithm was implemented using the DTA concept introduced in section 2.5.1.4. Figure 4.2 schematically describes the workflow of the DTD implementation: from the original dose distribution (A) two are created, B^+ and B^- (that correspond to A$\pm\delta$), and the DTA calculated. Then the DTD is the minimum value of the (DTA^+, DTA^-). The DTD in different regions of interest was compared to the variability in mapping between different algorithms, by measuring voxel-by-voxel the RMS of the DTD (DTD_{RMS}) and the Euclidean distance between the backward DVFs (also computed as a L^2-norm).

4.3 Results

4.3.1 Geometric matching

In Table 4.2 the values of DSC, OI, FN, FP, DT and CoM obtained averaged over both the directions of the registrations and for all structures/patients are presented. All the implementations provided similar results in terms of an overall geometric matching. The differences in DSC per structure were in general statistically insignificant ($p \in [0.1, 1]$) and of little clinical relevance. The two exceptions were an underperformance of DIR_{ics} in the

Table 4.2: *Mean values ± standard deviation of dice similarity index (DSC), overlap index (OI), false negative (FN), false positives (FN), distance transform (DT) and centroid position error (CoM). The results were averaged over all structures, patients and both registration directions.*

	$DIR_{sas+sas}$	$DIR_{sas+inv}$	DIR_{ics}	DIR_{svf}
DSC	0.851±0.080	0.847±0.082	0.848±0.075	0.851±0.073
OI	0.864±0.086	0.848±0.097	0.865±0.079	0.852±0.086
FP	0.17±0.11	0.15±0.11	0.18±0.12	0.15±0.10
FN	0.14±0.09	0.15±0.10	0.14±0.08	0.15±0.09
DT_{mean} (mm)	0.3±0.4	0.4±0.4	0.2±0.4	0.3±0.4
DT_{std} (mm)	1.3±0.4	1.3±0.3	1.3±0.4	1.3±0.3
$\|DT\|_{mean}$ (mm)	0.8±0.3	0.8±0.3	0.8±0.3	0.8±0.3
$\|DT\|_{std}$ (mm)	1.0±0.3	1.1±0.3	1.1±0.3	1.1±0.4
$\|DT\|_{95\%}$ (mm)	2.7±0.9	2.7±0.9	2.8±0.9	2.7±1.0
$\|DT\|_{max}$ (mm)	9±9	10±10	9±9	9±9
$\|DT\|_{2mm}$ (%)	9±6	10±6	10±6	9±6
CoM (mm)	1.1±0.9	1.2±1.0	1.3±1.2	1.2±0.9

external contours (p=0.02), and slightly better results for DIR_{svf} in the muscles (p<0.01). This was in agreement with the findings when visually inspecting the registrations: DIR_{ics} had difficulty in recovering larger and complex deformations, while DIR_{svf} performed particularly well in the alignment of soft tissues. The effect of inverse-consistency became evident when analysing the FN and FP results: FP≈FN for $DIR_{sas+inv}$ and DIR_{svf}, as the FP in one direction coincided with the FN in the other direction, and vice-versa, and so they averaged to similar values. FP≠FN for DIR_{ics} because even though this algorithm encourages inverse-consistency the resulting forward and backward transformations are only approximate inverses to each other.

4.3.2 Deformation field analysis

Table 4.3 shows the values of HE, det(Jac) and ICE found for each of the approaches. The voxels outside the patient were ignored when calculating the results. All the DVFs obtained had no negative det(Jac), thus were effectively invertible. In terms of physical plausibility of the DVFs, DIR_{svf} provided deformations with more desirable physical properties. Lower values of HE indicate smoother transformations, therefore the level of smoothness of the DVFs was higher in symmetric approaches (DIR_{ics} and DIR_{svf}), which resulted in tighter intervals of the det(Jac) values. Transformations constrained with an inverse-consistent penalty (DIR_{ics}) considerably improved the ICE in comparison with $DIR_{sas+sas}$, but the two resulting DVFs were clearly not real inverses as with $DIR_{sas+inv}$ and DIR_{svf}, for which the mean and standard deviation of the ICE were sub-mm. Figure 4.4 shows quantitatively the ICE for different registrations and its effect when composing

Table 4.3: *Mean values ± standard deviation for properties of the deformation vector fields: harmonic energy (HE), 1% and 99% percentiles of det(Jac), and mean, standard deviation and 99% percentile of the ICE. The results are averaged over all patients and both registration directions.*

	$DIR_{sas+sas}$	$DIR_{sas+inv}$	DIR_{ics}	DIR_{svf}
HE	0.14±0.05	0.16±0.05	0.11±0.05	0.11±0.04
$det(Jac)_{1\%}$	0.53±0.12	0.57±0.12	0.69±0.11	0.67±0.08
$det(Jac)_{99\%}$	1.49±0.16	1.57±0.18	1.33±0.13	1.40±0.13
ICE_{mean} (mm)	1.6±0.7	0.012±0.005	0.5±0.2	0.008±0.003
ICE_{std} (mm)	2.4±1.1	0.04±0.03	0.8±0.4	0.02±0.02
$ICE_{99\%}$ (mm)	13±6	0.12±0.07	4±2	0.06±0.03

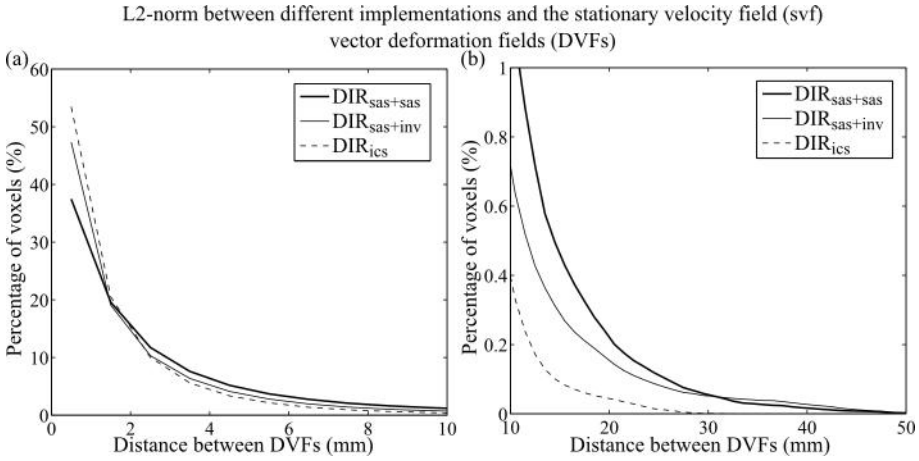

Figure 4.3: *L^2-norm between the deformation vector fields obtained with the stationary velocity field implementation and other algorithms.*

the forward and backward transformations on a slice of one of the patient's anatomy. $DIR_{sas+sas}$ and DIR_{ics} largest ICE values were found close to airways and in the shoulders region, where the CBCTs showed reduced contrast and higher noise. $DIR_{sas+inv}$ and DIR_{svf} maximum values of ICE resulted from numerical errors when composing the DVFs in regions of high contrast boundaries.

L^2-norm was calculated between DIR_{svf} and other approaches DVFs. Figure 4.3 plots the distribution of L^2-norm. The RMS value of the L^2-norm was 5±2 mm for $DIR_{sas+sas}$, 4.5±0.6 mm for $DIR_{sas+inv}$ and 2.6±1.3 mm for DIR_{ics}. DIR_{ics} produces DVFs more similar to DIR_{svf}, particularly when looking at the maximum L^2-norm values. However, 15% of voxels were mapped more than 2 mm apart by all methods, a distance clinically relevant.

4.3.3 Computation times

Table 4.4 shows the registration computation times measured for the different algorithms. When comparing the computation time taken to complete the registration in both

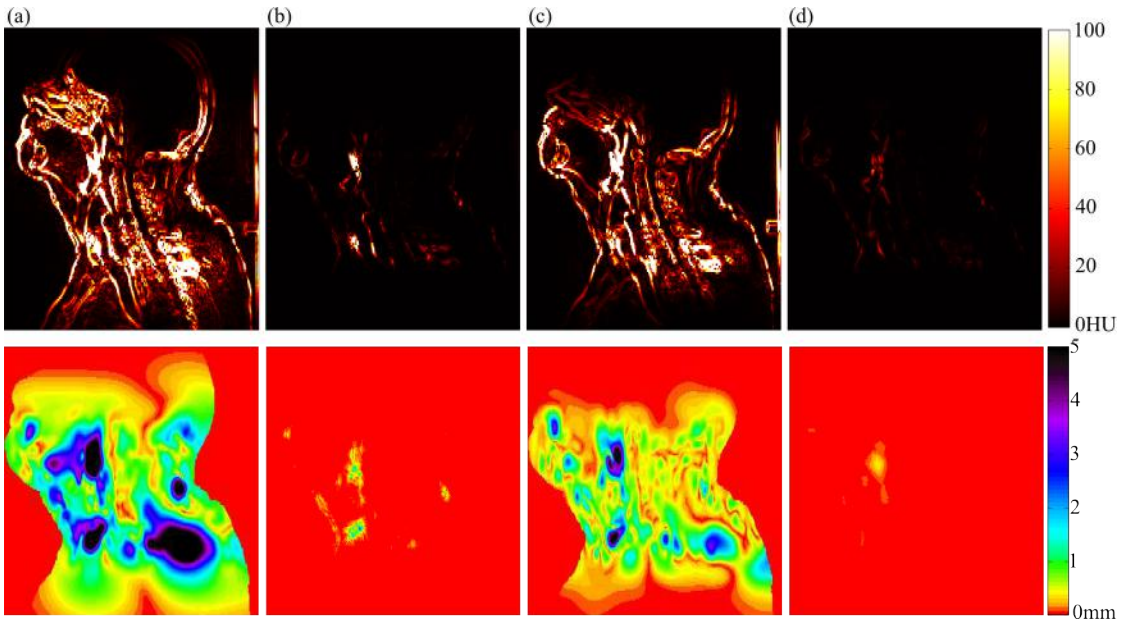

Figure 4.4: *Hounsfield unit variation after applying the composition of the forward and backward transformations to the pCT (top row), and inverse-consistency error (ICE) maps (bottom row) for a patient included in this study: (a) $DIR_{sas+sas}$, (b) $DIR_{sas+inv}$ (c) DIR_{ics}, and (d) DIR_{svf}.*

Table 4.4: *Mean values ± standard deviation registration of the computation times (in minutes) per DIR algorithm.*

	Forward	Backward	Total
$DIR_{sas+sas}$	15±3	8±5	23±6
$DIR_{sas+inv}$	15±3	5±1	20±3
DIR_{ics}		37±12	37±12
DIR_{svf}		61±15	61±15

directions and in the same processor, $DIR_{sas+sas}$ and $DIR_{sas+inv}$ resulted in similar times, while DIR_{ics} and DIR_{svf} took on average 2-3 times longer. In comparison, the standard forward asymmetric registrations ran in 0.9±0.2 min (range: 0.7-1.2 min) in the GPU. There are plans to implement all of the DIR algorithms available in NiftyReg so they can run on a GPU, but until then the use of DIR_{ics} and DIR_{svf} is limited to offline studies as current non-GPU computation times are too slow for online applications.

4.3.4 Dose warping comparison

Differences between the DVFs generated by different DIRs algorithms will affect the final warped dose distribution. Each of the two DVFs generated per patient (forward and backward) contributed to the differences in the final warped dose. First, differences in forward DVFs resulted in different dCTs and therefore different "doses of the day". Second, the backward DVFs remapped differently the "doses of the day" to the pCT

Table 4.5: *Mean values ± standard deviation for the dose difference test pass-percentage ($DD_{2\%\text{-}pp}$), mean value (DD_{mean}), root mean square value (DD_{RMS}), and 99th percentile of the DD distribution ($DD_{99\%}$) between different approaches and DIR_{svf} within different regions of interest (as a percentage of the prescribed dose (pD). The CBCT region stands for the imaging volume (dose voxels outside the patient were ignored). The treated volume (TV) corresponds to the volume encompassed by the planning 95% isodose surface, while the irradiated volume (IV) to the corresponds to the volume encompassed by the planning 50% isodose surface. Therefore IV-TV is the volume where 50 to 95% of the dose was planned to be delivered. P_{IV} was defined as the body slices close to the shoulders (a larger imaging volume) intersected with IV, where the CBCT HU were not reliable.*

	Method	CBCT FoV	TV	IV-TV	P_{IV}
$DD_{2\%\text{-}pp}$ (%)	$DIR_{sas+sas}$	81±5	91±3	79±5	72±7
	$DIR_{sas+inv}$	80±8	92±4	79±6	70±11
	DIR_{ics}	84±6	91±5	80±7	73±10
DD_{mean} (%pD)	$DIR_{sas+sas}$	1.6±0.4	0.8±0.2	1.7±0.3	2.1±0.4
	$DIR_{sas+inv}$	1.6±0.6	0.8±0.3	1.6±0.4	2.3±0.8
	DIR_{ics}	1.3±0.5	0.9±0.4	1.5±0.4	2.0±0.9
DD_{RMS} (%pD)	$DIR_{sas+sas}$	3.7±1.4	1.7±0.5	3.9±1.1	4.2±1.3
	$DIR_{sas+inv}$	3.2±1.3	1.6±0.7	3.1±0.8	4.2±1.6
	DIR_{ics}	2.9±0.9	1.9±0.8	3.3±0.9	3.9±1.6
$DD_{99\%}$ (%pD)	$DIR_{sas+sas}$	16±7	7±3	14±4	16±4
	$DIR_{sas+inv}$	15±6	7±3	14±4	19±7
	DIR_{ics}	11±4	9±5	13±4	15±8

space. The contribution of the first factor was small as the "doses of the day" differed by less than 2% of the prescribed dose on over 95% of the body voxels.

Dosimetric differences between the warped doses obtained with DIR_{svf} and every other approach were assessed. DIR_{svf} was arbitrarily chosen as the basis of this comparison since it generated the DVFs with more desirable physical properties.

Table 4.5 presents the DD found between different methods and DIR_{svf} in different regions of interest. There was no statistical evidence of any method being more similar to DIR_{svf} ($p \in [0.6\text{-}1.0]$). The differences were smaller in the treated volume (TV) and larger in the 50-95% of the prescribed dose volume (i.e., the irradiated volume (IV) that is not within the treatment volume, IV-TV), where higher gradients are more likely to occur. Regions of poorer CBCT quality (low contrast and high noise within larger imaging volumes [9], i.e., near the shoulders) within IV (P_{IV}) showed higher variability between warped doses. The differences between $DD_{2\%\text{-}pp}$ between all the different identified regions were statistically significant ($p<0.01$). Therefore regions of higher dose gradient and poorer CBCT image quality were more prone to having larger variability in warped doses, but for different reasons. DTD_{RMS} was 1.6±0.4 and 2.7±0.8 mm within IV-TV and TV, respectively. The RMS value of the L^2-norm within IV-TV, TV and P_{IV} was 2.7±1.1, 3.4±1.5 and 4.2±2.1 mm. L^2-norm values found for P_{IV} were statistically different from other regions ($p<0.01$),

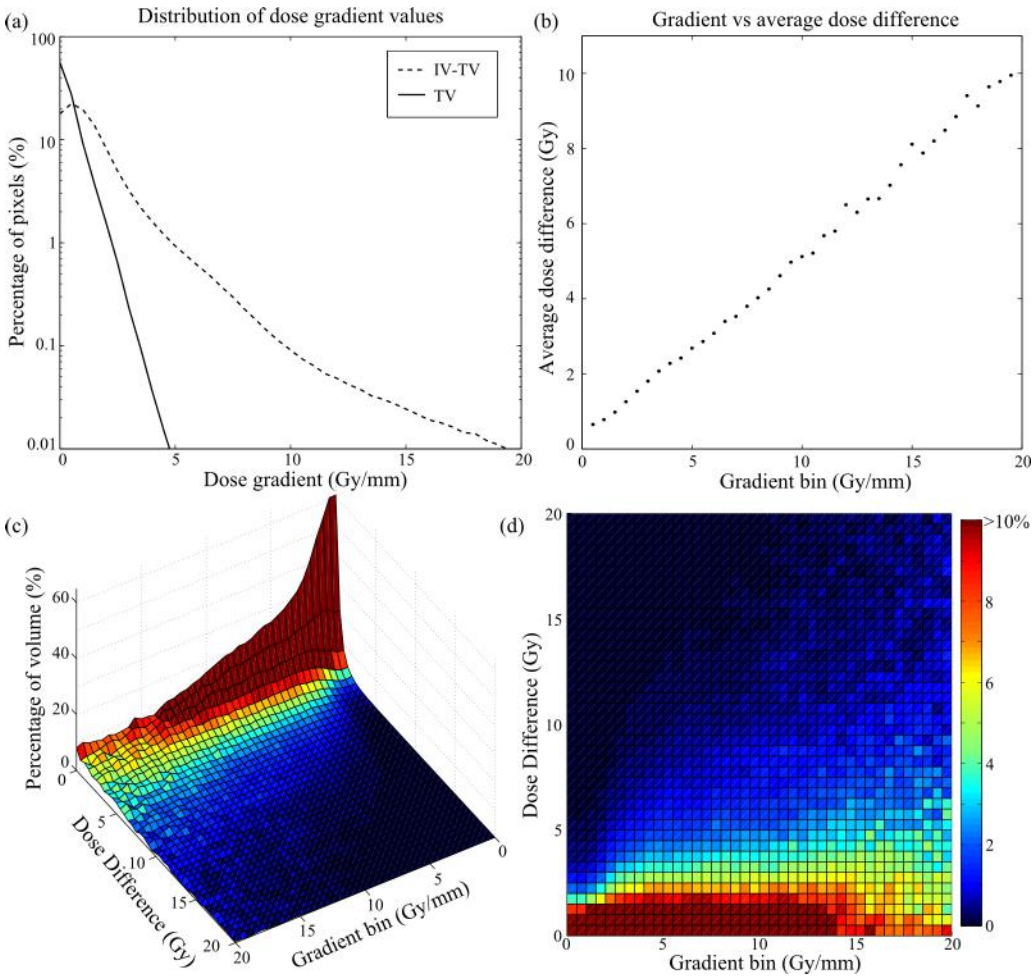

Figure 4.5: *Relationship between dose gradient and dose differences. (a) Distribution of dose gradient values within TV and IV-TV; (b) average dose difference as function of the dose gradient, and fraction of dose difference values per bin of dose gradient (c) 3D rendering and (d) top view.*

while between TV and IV-TV were not (p=0.1). Voxels within IV-TV had larger DD than voxels within TV due to the local characteristics of the dose distribution (shown by DTD), while inside P_{IV} the larger spatial mapping variability between DIR algorithms explained the larger DD (shown by L^2-norm).

The effect of the DVF in the dose mapping is complex, and was theoretically expected to depend on the location of the dose gradients [120, 124]. The dose gradient of the planned dose distribution was correlated with the values of DD within IV-TV. The correlation between gradient and DD was weak (Pearson correlation coefficient, ρ=0.281), but it was clear that as the dose gradient increased the distribution of DD values became more spread and the average DD increased (Figure 4.5).

The DD, ΔD_{mean} and ΔD_{max} to the spinal canal, brainstem, and parotids were computed (Table 4.6). The wide distribution of values found for ΔD_{mean} and ΔD_{max} was indicative that larger dose differences occur depending on the particular dose distribution and relative positioning of the OARs. When analysing the DVHs curves obtained for each patient it was found that in general all algorithms led to similar clinical outcomes, i.e.,

Table 4.6: *Mean values ± standard deviation (range) for the dose difference test pass-percentage ($DD_{2\%\text{-}pp}$), mean and root mean square of the DD (DD_{mean} and DD_{RMS}), and differences predicting the mean (ΔD_{mean}) and maximum doses (ΔD_{max}) to the OARs: spinal canal (SC), brainstem (BS), and parotids (PAR).*

	Method	$DD_{2\%\text{-}pp}$ (%)	DD_{mean}	DD_{RMS} (%pD)	ΔD_{mean} (%pD)	ΔD_{max} (%pD)
SC	$DIR_{sas+sas}$	90±12 (69-100)	0.8±0.5 (0.2-1.5)	1.1±0.6 (0.3-1.9)	0.4±0.6 (0.1-1.5)	1.0±0.9 (0.4-2.6)
	$DIR_{sas+inv}$	94±5 (86-100)	0.8±0.6 (0.2-1.9)	1.7±1.9 (0.3-5.0)	0.4±0.6 (0.0-1.4)	0.7±0.8 (0.0-2.0)
	DIR_{ics}	90±8 (77-96)	0.9±0.5 (0.5-1.7)	1.7±1.4 (0.8-4.2)	0.5±0.4 (0.2-1.1)	0.3±0.2 (0.0-0.5)
BS	$DIR_{sas+sas}$	92±14 (67-100)	0.8±1.0 (0.1-2.7)	1.4±1.9 (0.1-4.7)	0.6±0.8 (0.0-2.0)	0.6±0.4 (0.2-1.1)
	$DIR_{sas+inv}$	92±11 (72-100)	0.8±0.9 (0.1-2.3)	1.3±1.6 (0.1-4.2)	0.6±0.8 (0.0-2.0)	0.7±1.0 (0.0-2.5)
	DIR_{ics}	91±10 (79-100)	0.6±0.5 (0.1-1.2)	1.0±0.8 (0.2-2.0)	0.5±0.4 (0.1-1.1)	1.1±1.1 (0.2-2.8)
PAR	$DIR_{sas+sas}$	86±4 (79-93)	1.0±0.2 (0.6-1.3)	1.5±0.3 (1.1-2.0)	0.6±0.3 (0.2-1.0)	0.2±0.2 (0.0-0.5)
	$DIR_{sas+inv}$	86±7 (74-95)	1.0±0.3 (0.7-1.4)	1.5±0.4 (1.0-2.0)	0.3±0.4 (0.0-1.4)	0.3±0.3 (0.0-1.1)
	DIR_{ics}	83±9 (70-98)	1.1±0.5 (0.4-1.7)	1.9±0.8 (0.7-3.0)	0.9±0.5 (0.1-1.6)	0.2±0.2 (0.0-0.7)

that a replan was needed since OARs were receiving more dose than tolerated and targets less dose than planned. However, for one of the patients included in this study while $DIR_{sas+sas}$ and DIR_{ics} estimated a dose above the spinal canal tolerance, $DIR_{sas+inv}$ and DIR_{svf} did not (Figure 4.6). This is an example where the decision to replan could be affected by the choice of algorithm.

4.4 Discussion

Regarding the geometric evaluation, all the methods resulted in good alignment between anatomical contours. In some cases DIR_{ics} had worse results than the other approaches, indicating a reduced ability to capture more complex deformations due to the introduction of additional constraint terms. The differences in geometrical alignment from $DIR_{sas+sas}$ and DIR_{svf} were statistically and clinically insignificant, but the underlying properties of the deformations were different. DIR_{svf} resulted in deformations with more desirable physical properties, where both symmetry and inverse-consistency were satisfied.

Different DIR algorithms generated different DVFs which resulted in differences when warping the dose to the planning geometry. DIR_{ics} DVFs were found to be an overall better match to DIR_{svf}, but this did not result in higher similarity in warped doses. This could be explained by the fact that all algorithms had similar variability in mapping for differences larger than a clinical tolerance for distances (i.e., 2 mm). Also properties of the dose map will also have an effect in the dose uncertainties, not only the variability in mapping. The mean and maximum values found for DD between different DIR implementations were comparable to the values found by Salguero et al. when estimating the dose uncertainties of a DIR algorithm due to lack of inverse-consistency [123]. In this small feasibility study situations where the choice of algorithm led to higher uncertainties in dose warping were

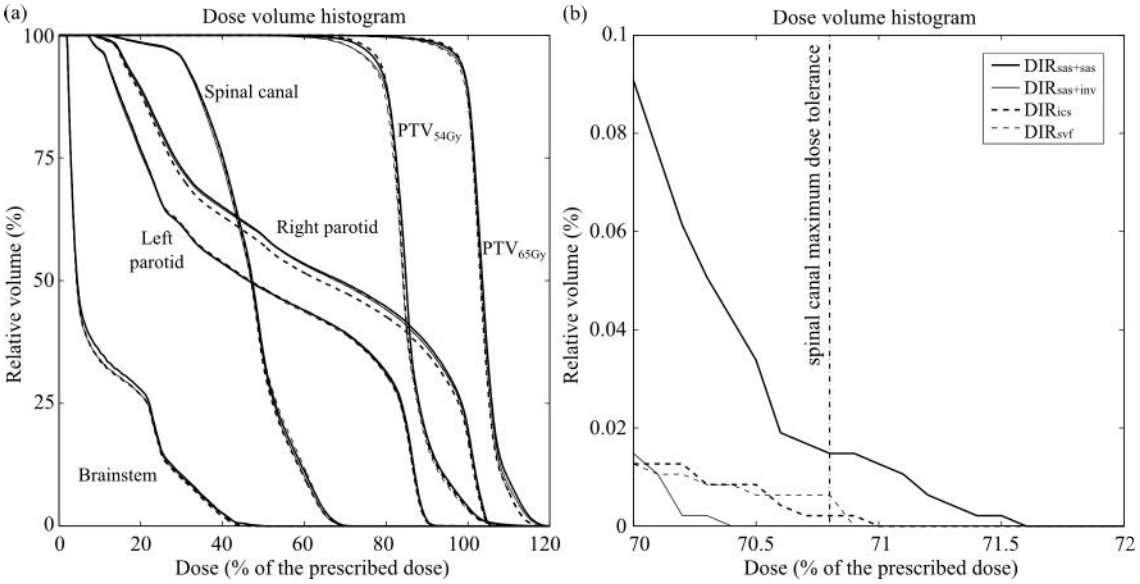

Figure 4.6: *Dose volume histogram obtained for a patient included in this study, using doses warped by different DIR algorithms: (a) all organs-at-risk and target volumes and (b) zoom of the maximum dose to the spinal canal. $DIR_{sas+sas}$ and DIR_{ics} estimated a dose above the tolerance, while $DIR_{sas+inv}$ and DIR_{svf} did not.*

identified. The first important point was the effect of the dose gradient. Where there was a high dose gradient there was more likely a larger variability in dose between the different DIR algorithms. OARs within the high gradient region can be of concern, as different methods could predict maximum doses to the spinal canal and brainstem with a difference of up to 2.8%pD. The correlation between gradient and DD was weak and similar to the values reported in the same study by Salguero et al.. The weak correlation can be explained due to the fact that if a registration error occurs in uniform dose region the resulting dose error will be small, but when a registration error occurs in high dose gradient the resulting dose difference may be large but does not mean it will be, since there are other factors to consider besides the gradient (i.e, how large the uncertainty in spatial mapping is, whether the registration error is in the same direction as the gradient). The results obtained are also in agreement with the findings of Saleh-Sayah et al., as the required spacial accuracy depended on the local dose distribution [120]. A second important point is the effect of the lower quality of the CBCT in the registration uncertainty. Regions of reduced contrast and increased noise (particularly evident in larger imaging volumes like the shoulders) were more susceptible to variability in mapping between registration algorithms, and therefore in larger differences between warped doses. The impact of the choice of DVF will depend on the dose distribution and relative positioning of OAR, and generic validation frameworks (based exclusively on geometric analysis of the deformations as explored in chapter 3) are not sufficient for dose warping applications. The patients included in this study were a challenging cohort to test DIR, and it is likely that the issues reported may have less impact for less demanding patients (but then these patients will benefit less from ART). A larger study could potentially identify patterns

of when different algorithms provide significant differences between doses, and flag the regions where the dose warping algorithm may be unreliable.

For an application that is sensitive to the underlying deformations such as dose warping we believe a more complex algorithm like DIR_{svf} is preferable over other approaches. This opinion is based on the similar ability to recover deformations while generating deformations with more desirable physical properties. Other studies support the theoretical advantages of ensuring symmetry and inverse-consistency to improve the precision of dose warping using DIR. Bender et al. studied the effect of inverse-consistency and transitivity in DIR for a single HN patient [39]. Lack of Transitivity means that different dose distributions will be obtained depending on the order in which the registrations are used, and the time point chosen for summation. They found dose differences at OAR when different image time points were used as a reference for summation; however, when increasing the inverse-consistency and transitivity those differences were considerably reduced. Inverse-consistent algorithms do not enforce transitivity making this an interesting area of research for registration developers. Yan et al. showed that dose mapping ICE observed when mapping doses back-and-forth was reduced 1.5-3 times when the spatial inverse-consistency was improved [38].

The reliability of using dose warping in clinical settings is a current and open debate [130]. This work contributes to this discussion by evaluating theoretically better DIR algorithms and investigating the uncertainties in dose warping due to the choice of algorithm in ART frameworks that use CT-to-CBCT registration. One of the main difficulties with validating dose warping is that the true point to point mapping is difficult or impossible to establish, especially in regions of homogeneous image intensities as is often found inside individual structures [100, 131]. When new plans are based on accumulated dose, registration inaccuracies will also affect the newly planned treatment. The difficulties inherent to validate DIR for dose warping do not necessarily discourage its use for clinical research, but users should carefully consider their choice of DIR algorithm and the conclusions that should be drawn from the results. Extensive and detailed studies on the behaviour of DIR algorithms can be particularly problematic when using commercial products that, unlike NiftyReg, do not allow flexibility in selecting appropriate parameters [130, 132].

It should be noted that the real changes occurring in the tissue are complex and variate: sometimes tissue appears or disappears (e.g. weight loss and tumour shrinkage) and sometimes it expands/contracts or deforms in other ways. The vast majority of current DIR algorithms use a transformation model that represents expansion/contraction, but map constant image intensities, which in CT represents a constant density and so is more representative of appearing/disappearing tissue. To accurately model and recover the real physical changes that occur during a course of radiotherapy is extremely challenging but it is what an ideal DIR algorithm should be able to do and what should be an aim of the next generation of DIR algorithms. Several groups are actively working on making DIR algorithms more realistic. Examples include incorporating missing tissue in

image registration by modifying existent DIR algorithms [133] and further regularising the transformation to avoid deformation of bony anatomy [110]. However, there is still a very long way to go to achieve truly realistic DIR, and indeed this will not just involve developing new algorithms and computational techniques, but will also require a better understanding of the actual physical and biological processes that occur during a course of radiotherapy.

4.5 Conclusions

This chapters presents the evaluation of several DIR algorithms for CT-to-CBCT registrations and investigated the uncertainties inherent to using different DIR algorithms to warp doses to a reference geometry. Standard asymmetric and stationary vector field implementations resulted in similar geometric matching, but the properties of the DVFs were very different, with the second providing deformations with more desirable physical properties. The choice of DIR implementation had a larger impact on the dose warped in regions where the dose gradient was higher and/or the CBCT image quality was poorer. The reliability of using dose warping and summation in clinical settings is indeed a topic of interest; however, the lack of gold-standards for validation of DIR is still an unsolved issue. Further understanding of the limitations of current DIR algorithms is necessary before clinical translation.

Chapter 5

Head and neck proton adaptive therapy

> If I have seen further it is by standing on the shoulders of giants.
>
> Isaac Newton

This chapter consists of an extension of the work reported in chapter 3 to proton therapy. The appropriateness of the translation of the CBCT and DIR method to this treatment modality was investigated and compared to the results previously obtained for photon treatment. The most recent registration algorithms (used in chapter 4) and similarity measures available in NiftyReg were utilised for this clinical application.

The work in this chapter resulted in the following outputs:

- C. Veiga, J. Alshaikhi, R. Amos, A. M. Lourenço, M. Modat, S. Ourselin, G. Royle, and J. R. McClelland, "CBCT and deformable registration based "dose of the day" calculations for adaptive proton therapy," Int. J. Particle Ther. 2(2) 404-414 (2015).
- C. Veiga, J. Alshaikhi, M. Modat, S. Ourselin, G. Royle, R. Amos, and J. R. McClelland, "CBCT and deformable registration based dose calculations for adaptive proton radiotherapy," 4D Treatment Planning Workshop (London, United Kingdom, 2014).

5.1 An introduction to proton therapy

Proton therapy optimises radiation treatment by delivering radiation doses with great precision due to the finite range of protons within matter [4]. The different shape of the dose-depth curves obtained using different types of particles (Figure 5.1), and therefore

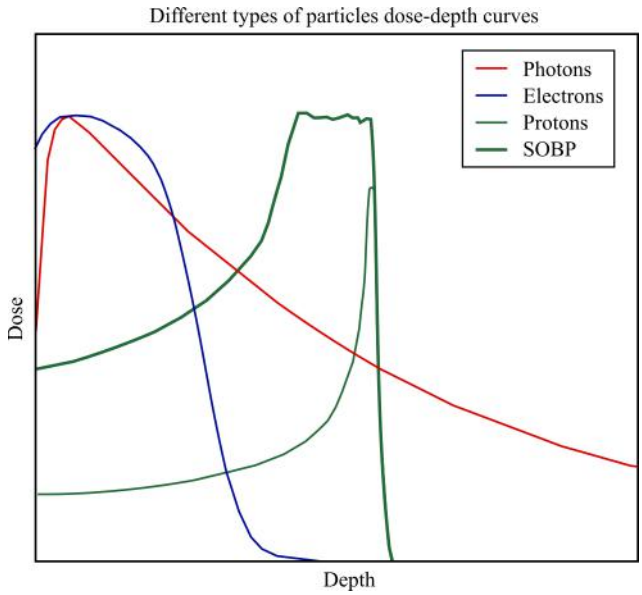

Figure 5.1: *Schematic dose-depth curves of different particles: photons, electrons and protons.*

precision of cancer treatment with such particles, is related with the physical processes that occur when interacting with matter [134]:

1. Photons are electrically neutral and interact stochastically with matter. Photons do not steadily lose energy as they penetrate matter, but instead travel until absorbed or scattered (i.e., changing direction of travel, with or without loss of energy). The probability of interaction depends on its energy and medium traversed, and is known as the linear attenuation coefficient. Therefore, photons do not have a finite penetration depth, and several physical processes can remove photons from the primary beam: photoelectric effect, Compton effect, pair production, photo-nuclear reactions and Rayleigh scattering.

2. A proton traversing matter loses energy primarily by ionising and exciting the atoms of the medium. A heavy charged particle is only able to transfer a small percentage of its energy in a single collision with an atomic electron, and its deflection is negligible. Therefore, a proton travels an almost straight path through matter (they are slightly deviated from their path due to multiple Coulomb scattering), losing energy almost continuously in small amounts. Charged particles have a finite penetration depth unlike photons, and exhibit a low ionisation density at surface, that slowly increases till almost the end of their range, at which point there is a narrow peak of high ionisation density (known as Bragg Peak), with negligible dose deposition after this point. The dose falls off sharply both laterally and distally [135]. Due to the statistical nature of the interactions, all particles with the same initial energy do not travel exactly the same distance. This phenomenon is called range straggling. The average linear rate of energy loss of a heavy charged particle in a medium, –dE/dx, is called the stopping power.

3. Electrons are also charged particles and thus lose energy almost continuously as they slow down in matter. In contrast to protons, electrons do not generally travel in straight lines; because of their smaller mass they can lose a large fraction of their energy in a single electronic collision, and are frequently scattered large angles by nuclei. In addition, they can also radiate energy by bremsstrahlung. Electron beams are therefore characterised by a maximum range in tissue beyond which one can see a low intensity tail due to the bremsstrahlung photons.

A pristine Bragg peak cannot be used in cancer treatment as it is too narrow when compared with the normal size of a tumour. The solution is to irradiate the tumour volume uniformly by combining proton beams of different energies into a spread-out Bragg peak (SOBP) [136]. To produce a SOBP of desired width and shape the beam to the tumour, two delivery techniques are available in proton therapy gantries: scattering and scanning systems (Figure 5.2). In a passive scattering proton therapy (PSPT) the primary proton beam transverses a scattering system to broaden the beam to treatment dimensions, and the energy is degraded using range modulator wheels. Patient specific hardware, like blocks and compensators, are used to shape the beam according to the target volume. An alternative to scattering systems is to scan a monoenergetic pencil beam magnetically across the target volume. Different depths are scanned using different energies and no beam shaping devices are needed. Scanning systems are commonly used in clinical settings to create single-field uniform dose (SFUD) plans, where each beam spot positions and weights are optimised independently, such that each beam covers uniformly the target. However, scanning system also allow true intensity modulated proton therapy (IMPT), where the beam spots from multiple beams are optimised simultaneously (this is also often called multiple-field optimisation (MFO)).

5.2 Rationale

The introduction of IMRT as state of the art treatment considerably improved the survival and quality of life of HN patients [137]. However, achieving optimal conformality to target volumes with IMRT can lead to substantial irradiation of the brainstem, oral cavity, salivary glangs, cochlea, larynx and optic apparatus, which can cause both acute and chronic morbidity. The reduction in integral dose made possible with the development of proton therapy may result in improved care of HN patients [138].

Numerous reports have been published documenting the theoretical advantages [138–141] and clinical evidence [142–146] of the benefits of proton over photon therapy for HN malignancies. The general outcomes of these studies show that disease free survival and overall survival were comparable between photon and proton therapy, while local control was statistically significantly increased. It is possible with proton therapy to achieve decreased doses to the optical nerves, to the parotid glands and to the oral cavity, leading to reduction in feeding tubes used. This is accomplished without compromising

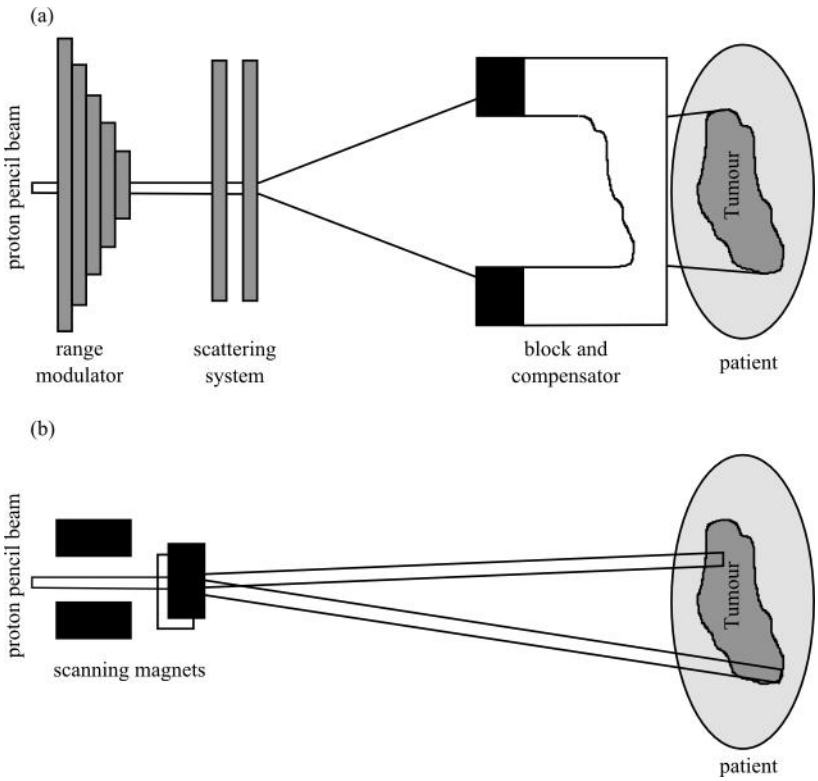

Figure 5.2: *Schematic of (a) passive scattering and (b) scanning proton delivery systems.*

the conformity index, dose homogeneity and coverage to the target volume [57].

It has been subject of debate whether the high initial and treatment costs of particle therapy are justified given the currently available evidence of its effectiveness. Based on the available evidence, particle therapy seems at least as effective as photon therapy in tumour control, and may have advantage in sparing OARs. The superiority in terms of tumour control and survival remains uncertain. However, the number of published studies is limited and the data gathered in less than optimal physics-based settings, making difficult to establish the real effectiveness of proton therapy in HN care [57]. The majority of clinical experience in HN is with a combination of traditional photon therapy and PSPT [144]. Currently, insufficient data is available to recommend proton therapy for routine HN cancer treatment outside of clinical trials [147]. University of Texas MD Anderson Cancer Center (Houston, TX, USA) and University of Pennsylvania Roberts Proton Therapy Center (Philadelphia, PA, USA) have recently started reporting the treatment of HN with spot scanning; nevertheless, it is within the plans of the UK proton center project to prioritise the treatment of HN malignancies.

Although there are a plethora of image-guidance techniques available in the photon world, the same technology is still underdeveloped when it comes to proton therapy. The fundamental physical differences between photon and proton interactions with the matter make photon imaging techniques sub-optimal for proton therapy. The most immediate example is CT. Stopping power of different tissues cannot be obtained directly from the electron density extrapolated from CT data for treatment planning. As consequence,

the current practice is to convert CT values to material composition and density using stoichiometric approaches [148]. Taking advantage of the properties of protons has been a major motivation to develop proton specific imaging modalities, such as proton radiography/CT [149] for treatment planning, and positron emission tomography (PET) [150] or prompt gamma emissions [151] for *in vivo* range verification. However, such imaging modalities are still in research or development stage and not available clinically, which has somehow stalled the translation of photon imaging modalities to the proton clinic over the last decade. Most operational centres worldwide only perform image-guidance with planar kV imaging, and only recently bidimensional (2D) real-time imaging and gating systems have been reported [152]. Translating existing 3D photon imaging technology, particularly CBCT, into the proton clinic has recently gained interest. The development of CBCT in proton therapy systems required solving engineering problems related with the geometry of a proton-gantry. Seabra et al. described the major challenges for CBCT integration in proton therapy gantry compared to a LINAC [153]:

- Source to isocenter distance is 3 times longer than in a LINAC, so a larger x-ray power is necessary. This requires a higher thermal capacity of the tube for a complete arc.
- Nozzle to isocenter distance is approximately half than in a LINAC. The smaller clearance imposes additional safety measures to avoid collision between parts.
- The flat panel detectors are fixed in tractable arms extending from the back of the gantry. The long cantilevered distance leads to deviation of the panel trajectory.

At the time of the conclusion of the work included in this chapter (November 2014), several groups were conducting the commissioning of CBCT systems for clinical use in proton therapy [153–155], in collaboration with manufacturers that already advertised the selling of such products with their proton therapy solutions. The world's first prototype of a clinical CBCT system on a proton-gantry was installed at Roberts Proton Therapy Center (Philadelphia, PA, USA) during the summer of 2014; however, patient data only started being acquired regularly from December 2014. In chapter 6 the first patient data acquired on a proton-gantry mounted CBCT will be used in the context of adaptive lung proton therapy. Currently, this technology is still not widely available, and in-room CT is the only reliable alternative for volumetric image-guidance. The clinical availability of in-room CT is however also still very limited [156].

Although theoretically protons have dosimetric advantages versus photons, to fully take advantage of the potential of proton therapy and achieve clinical benefit, the variations in anatomy (such as weight loss and tumour shrinkage) have to be monitored and accounted for, as it is becoming popular in photon therapy [73]. In chapter 3 the feasibility of using DIR to map the HU from the CT to the geometry of the CBCT was investigated. This method provided a good estimation of the "dose of the day" for IMRT treatments which can be used to feed an ART workflow. However, the challenges found in conventional photon therapy are even more concerning for proton therapy, and this

approximation may no longer be valid for proton therapy mainly for two reasons. First, since the dose gradient in proton dose distributions is steeper, accurate positioning is even more crucial to minimise the risk of overdosing OARs and/or underdosing target volumes. Additionally, fraction to fraction changes in size and position of tissue heterogeneities will adversely affect the dose distribution properties of protons to a greater extent than photons since proton dose-depth curves are more dependent on the physical properties of the tissue than photons. In the HN region the changes can be complex as protons travel through a complex anatomy composed of air, bone, and soft tissues [138]. Proton dose calculations are therefore expected to be more sensitive to registration errors than analogous IMRT cases. The aim of this work is to evaluate the feasibility of a CBCT and DIR based "dose of the day" calculation for adaptive proton therapy, which was previously evaluated for IMRT treatments.

5.3 Methods

5.3.1 Patient data acquisition

A total of three HN patients were used in this study. This selection corresponded to PT# 3, 4 and 5 described in Table 3.1 from the cohort described in sections 2.5.1.1 and 3.2.1.

5.3.2 Treatment planning

The choice of beams was to optimise target coverage while minimising dose to the brainstem, spinal canal, oral cavity, salivary glands and larynx. The same prescribed doses and volumes were used in all treatment approaches. 65 Gy(RBE) to the primary disease and 54 Gy(RBE) to the secondary disease were planned to be delivered in 30 fractions. The planning target volumes (PTVs) were defined as a 3 mm expansion of the clinical tumour volumes (CTVs). The primary objective of the plans was to achieve 95% of the prescribed dose to the PTV while maximising conformity. The plans were then optimised to minimise the dose to the OARs without compromising target coverage. Tolerance doses followed the UCLH guidelines for IMRT treatments: maximum doses of 46 Gy(RBE) and 55 Gy(RBE) to the spinal canal and brainstem, and mean unilateral dose of 20 Gy(RBE) and bilateral dose of 25 Gy(RBE) to the parotid glands.

Treatment planning was performed on the pCT scan using Eclipse External Beam planning system (version 10.8, Varian Medical Systems, Palo Alto, CA, USA). For IMRT cases the patients had been planned as part of UCLH's clinical workflow using a 7-field protocol, and were treated with the plans used in this study. Thus, the IMRT plans used were the same as chapter 4), but do not match those of chapter 3 (i.e., the initial plan was used here instead of the replan). Proton treatment planning was done retrospectively by

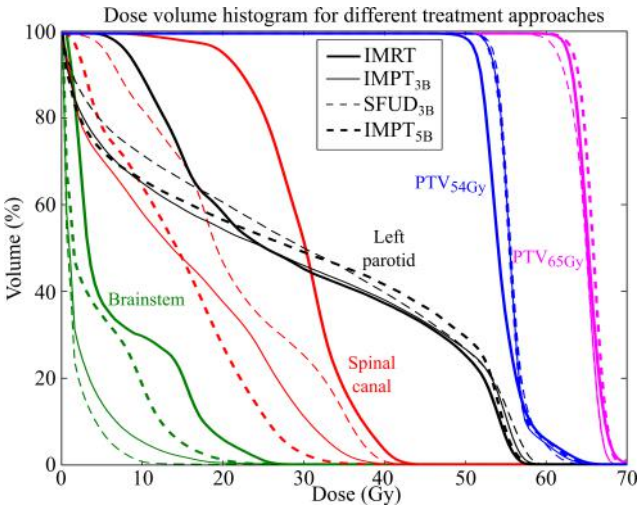

Figure 5.3: *Dose volume histogram comparing proton and photon plans. Right parotid omitted for figure clarity.*

Jailan Alshaikhi and Dr Richard Amos (Radiotherapy Physics, UCLH), with the plans generated per patient consisting of two types of optimisation: IMPT and SFUD, and two different beam arrangements: 3 beams with gantry rotations of 60°, 180° and 300° and 5 beams with gantry rotations of 45°, 135°, 180°, 225°, and 315°. A total of three plans were generated, IMPT with 3 and 5 beams ($IMPT_{3B}$ and $IMPT_{5B}$), and SFUD with 3 beams ($SFUD_{3B}$). $IMPT_{3B}$ represents a standard curative approach, which maximises the potential benefits of proton therapy (i.e., reduced integral dose, minimised dose to OARs and higher homogeneity inside the PTV) [141, 146]. However, such plans can be sensitive to positioning errors and anatomical changes. In contrast, $IMPT_{5B}$ and $SFUD_{3B}$ are more robust planning strategies at the cost of smaller dosimetric benefits. Particularly, $SFUD_{3B}$ reduced the ability to minimise the maximum dose to the spinal canal as all the fields have similar weight, while $IMPT_{5B}$ increased the integral dose. Table 5.1 presents the dose statistics and properties of the plans, such as mean and maximum doses to OARs, conformity index [157], homogeneity index [158] and non-target integral dose [159]. The focus of this study was not treatment planning, and therefore the proton plans were designed to be clinically acceptable and to demonstrate the benefits of proton therapy, and not necessarily to be optimal (Figure 5.3).

5.3.3 Image registration settings

The implementation based on the stationary velocity field transformation model available in NiftyReg [37] was used. The DIR algorithm does not match that of chapter 3, and a different measure of similarity was used here in comparison to chapter 4. Instead of NMI, LNCC was used as similarity measure as it is well suited to account for the differences in image intensities between CT and CBCT. As LNCC calculates the similarity over local windows rather than the whole volume, it can better account for the spatially varying intensity values in the CBCT image than global similarity measures such as NCC or SSD.

Table 5.1: *Mean ± standard deviation of the dose statistics and properties of the plans used in this study.*

		IMRT	$IMPT_{3B}$	$SFUD_{3B}$	$IMPT_{5B}$
Spinal	D_{mean} [Gy(RBE)]	28±4	13±2	20.5±1.4	11±3
canal	D_{max} [Gy(RBE)]	41±2	33±4	42±2	28±3
Brainstem	D_{mean} [Gy(RBE)]	14±6	87±4	7±4	5.6±1.0
	D_{max} [Gy(RBE)]	38±9	29±9	24±11	25±4
Left	D_{mean} [Gy(RBE)]	40±7	29±5	34±5	29±4
parotid	D_{max} [Gy(RBE)]	63±5	63±4	64±4	64±5
Right	D_{mean} [Gy(RBE)]	43±6	33±2	37.8±0.7	33.3±1.6
parotid	D_{max} [Gy(RBE)]	67.1±1.6	66.8±0.4	67.1±0.1	66.9±0.4
Conformity Index[a]		0.87±0.02	0.81±0.06	0.88±0.03	0.80±0.05
Homogeneity Index[b]		1.08±0.02	1.08±0.02	1.10±0.02	1.10±0.07
Integral Dose[c] [Gy(RBE)×L]		160±40	120±20	120±20	130±20

[a]$CI=V_{PTV} \cap V_{95\%}/V_{95\%}$, where V_{PTV} is the volume of the PTV and $V_{95\%}$ the volume of the 95% isodose level.

[a]$HI=D_{5\%}/D_{95\%}$, where $D_{5\%}$ and $D_{95\%}$ are doses received by 5% and 95% of the PTV

[a]$ID=V_{NTV} \times D_{mean}$, where V_{NTV} is the non-target volume and D_{mean} its mean dose.

The out-of-field approximation was reliable for IMRT treatments, but mostly in the superior direction, where the anatomy moved rigidly. The limited FoV is expected to be more problematic in proton therapy so the suitability of this approximation is evaluated in this work.

5.3.3.1 Geometric matching and properties of the deformation fields

The registrations using the LNCC and NMI were compared in terms of the following using the following quantitaties: DSC, OI, FN, FP, DT and CoM (section 2.5.1.3 and 3.2.3.1); HE, properties of the determinant of the Jacobian [det(Jac)], and ICE. The structure set consisting of the external contours, vertebrae C1,C4 and C7, RSCM and LSCM (section 2.5.1.3). The full cohort of five HN patients (versus three in the rest of this chapter) was used in this part of the study.

5.3.4 Dose comparison

DIR was used to map the HUs from the pCT to a CBCT that closely resembled the rCT (the calibrated dCBCT described in section 3.2.3.2). A dosimetric evaluation was performed to evaluate the impact of the registrations errors in calculating the "dose of the day". The data and registrations used match those of Figure 3.5. Dose distributions recalculated on rCT were considered gold-standard, and compared to doses on dCT (our

method) and directly on a calibrated dCBCT and rigidly-aligned pCT (alternative methods). Therefore, dose distributions for the same IMRT and proton plans were calculated on the pCT, rCT, dCT and calibrated dCBCT (D_{pCT}, D_{rCT}, D_{dCT} and D_{CBCT} respectively). The isocenter was placed in the same point in both cases based on rigid alignment of the vertebrae. All doses were calculated with a resolution of 2 mm.

The uncertainty of the dose calculations was evaluated by computing the voxelwise difference between dose distributions, known as the DD test, and using dose-volume histograms DVHs, similarly to the methodology used in section 3.2.3.2.

5.4 Results

5.4.1 Geometric validation

Table 5.2 presents the results of the performance of LNCC in the geometrical and deformation field analysis, in comparison with the results from using NMI in chapter 4. The geometrical accuracy is considerably improved when using LNCC, while the underlying global properties of the deformation are relatively unchanged. Particularly, there is an increase in the values of DSC and OI, and the mean value of the signed DT becomes closer to zero.

5.4.2 Dose comparison

Tables 5.3 and 5.4 provides the results obtained for $DD_{2\%-pp}$ and DD_{RMS} between the different methods and the rCT within different regions of interest. Due to the anatomical changes, the pCT does not give a good estimate of the "dose of the day", particularly for proton plans. Dose calculations on the calibrated dCBCT also result in a poor estimation of the "dose of the day" in spite of the images closely representing the anatomy of the day. Visual inspection of the DD maps revealed that the major source of dose mismatch was within regions where the CBCT imaging quality was degraded (such as nearby the shoulders where the imaging volume is larger). In such larger volumes a single calibration curve failed to recover the correct HU, resulting in inaccurate dose calculations. Proton plans were therefore more sensitive both to anatomical changes and to the inconsistency in HU characteristic of CBCT imaging than photon plans. The errors in the "dose of the day" calculations on the dCT were also larger for the proton plans than for the IMRT plans, but in all cases were considerably lower than for the calibrated CBCT and rigidly-aligned pCT.

Different regions of the dose map were more sensitive to registration errors. The $DD_{2\%-pp}$ differences between the volume encompassed by the planning 95% isodose surface (TV), and the volume receiving 50-95% of the prescribed dose (the IV minus

Table 5.2: *Mean values ± standard deviation of dice similarity index (DSC), overlap index (OI), false negative (FN), false positives (FN), distance transform (DT),centroid position error (CoM), harmonic energy (HE), properties of the determinant of the Jacobian [det(Jac)], and inverse consistency error (ICE) for NMI and LNCC registrations. The results are averaged for all patients, structures or DVFs, and registration directions.*

	DIR_{NMI}	DIR_{LNCC}
Geometric matching		
DSC	0.851±0.073	0.863±0.067
OI	0.852±0.086	0.863±0.076
FP	0.15±0.10	0.14±0.08
FN	0.15±0.09	0.14±0.08
DT_{mean} (mm)	0.3±0.4	0.1±0.4
DT_{std} (mm)	1.3±0.3	1.2±0.3
$\|DT\|_{mean}$ (mm)	0.8±0.3	0.7±0.2
$\|DT\|_{std}$ (mm)	1.1±0.4	1.0±0.3
$\|DT\|_{95\%}$ (mm)	2.7±1.0	2.5±0.8
$\|DT\|_{max}$ (mm)	9±9	9±10
$\|DT\|_{2mm}$ (%)	9±6	8±5
CoM (mm)	1.2±0.9	1.1±0.9
Characteristics of the deformation fields		
HE	0.11±0.04	0.13±0.02
$det(Jac)_{1\%}$	0.67±0.08	0.66±0.07
$det(Jac)_{99\%}$	1.40±0.13	1.43±0.12
ICE_{mean} (mm)	0.008±0.003	0.009±0.003
ICE_{std} (mm)	0.02±0.02	0.018±0.011
$ICE_{99\%}$ (mm)	0.06±0.03	0.07±0.03

Table 5.3: *Mean ± standard deviation of the dose difference test pass-percentage ($DD_{2\%-pp}$) between doses calculated using a rigidly-aligned planning CT, calibrated CBCT and deformed CT in comparison with a replan CT within different regions of interest. The CBCT and non-imaged field-of-view (FoV) stand for the regions where more than 10% of the prescribed dose was deposited that were imaged or not, respectively. The treated volume (TV) corresponds to the volume encompassed by the planning 95% isodose surface, while the irradiated volume (IV) corresponds to the volume encompassed by the planning 50% isodose surface. Therefore IV-TV is the volume where 50 to 95% of the dose was planned to be delivered.*

		Non-imaged FoV	CBCT FoV	TV	IV-TV
IMRT	D_{pCT}	92.3±1.2%	72±6%	64±16%	73±5%
	D_{CBCT}	89±4%	74±3%	84±9%	62±5%
	D_{dCT}	96.0±0.6%	93.2±0.7%	99.1±0.4%	94.3±0.2%
$IMPT_{3B}$	D_{pCT}	59±12%	51±4%	49±9%	32±8%
	D_{CBCT}	65±7%	62±3%	72±5%	44±2%
	D_{dCT}	76±6%	85±2%	88.8±0.3%	71.1±1.1%
$SFUD_{3B}$	D_{pCT}	67±8%	62±2%	81±8%	43±11%
	D_{CBCT}	69±7%	69.7±0.8%	91±4%	54±5%
	D_{dCT}	80±3%	87±2%	97.7±0.8%	76±5%
$IMPT_{5B}$	D_{pCT}	65±9%	57±8%	52±11%	36±7%
	D_{CBCT}	68±4%	66±3%	76±8%	44.7±0.8%
	D_{dCT}	79.3±0.4%	88±3%	91±4%	75±3%

the TV, IV-TV), where higher gradients are more likely to occur, indicate that the local properties of the dose map affect the accuracy of the "dose of the day" calculations. Proton plans were also more sensitive than photon plans to higher dose gradients. For example, for $IMPT_{3B}$ the DD_{RMS} was 2.6±0.6%pD within TV, and this value increased to 8.2±0.4%pD within IV-TV. Similar behaviour was found for all the plans, but in IMRT the DD_{RMS} was considerably lower (0.57±0.11%pD and 3.1±0.3%pD for TV and IV-TV, respectively). Figure 5.4 provides a qualitative view of this effect. While in IMRT most errors occurred near skin and airways, in proton plans the differences are larger within the high dose and dose gradient regions. Additionally, in proton plans the $DD_{2\%-pp}$ outside the imaging FoV was consistently smaller than within the CBCT FoV, unlike the IMRT case. This is indicative that while using the pCT outside the imaging FoV was a valid approximation for proton treatments (particularly in the superior direction) special care is needed if high dose gradients occur outside the imaging FoV.

Different optimisation and delivery techniques for the proton plans resulted in different $DD_{2\%-pp}$ and DD_{RMS}. $IMPT_{3B}$ is the approach more sensitive to DIR errors, particularly within IV-TV. $SFUD_{3B}$ performed particularly well within TV, since all fields deliver a uniform dose to the high dose region, making it less sensitive to DIR errors. In general $IMPT_{5B}$ also performed better than $IMPT_{3B}$ as less dose is delivered outside the target per beam. The DIR method appears to perform better in photon cases, while for proton cases

Table 5.4: *Mean ± standard deviation of the root mean square of the dose difference distribution (DD_{RMS}) between doses calculated using a rigidly-aligned planning CT, calibrated CBCT and deformed CT in comparison with a replan CT within different regions of interest. The CBCT and non-imaged field-of-view (FoV) stand for the regions where more than 10% of the prescribed dose was deposited that were imaged or not, respectively. The treated volume (TV) corresponds to the volume encompassed by the planning 95% isodose surface, while the irradiated volume (IV) corresponds to the volume encompassed by the planning 50% isodose surface. Therefore IV-TV is the volume where 50 to 95% of the dose was planned to be delivered.*

		Non-imaged FoV	CBCT FoV	TV	IV-TV
IMRT	D_{pCT}	2.5±0.2%pD	8.6±1.1%pD	4.6±1.6%pD	8±3%pD
	D_{CBCT}	2.1±0.2%pD	3.5±0.3%pD	1.8±1.6%pD	4.0±0.7%pD
	D_{dCT}	1.5±0.4%pD	2.5±0.2%pD	0.57±0.09%pD	3.1±0.3%pD
$IMPT_{3B}$	D_{pCT}	5.9±1.1%pD	10±3%pD	6.0±1.7%pD	14±4%pD
	D_{CBCT}	5.2±1.5%pD	6.7±0.9%pD	3.3±0.5%pD	9.9±1.5%pD
	D_{dCT}	3.3±0.8%pD	4.0±0.1%pD	2.6±0.6%pD	8.2±0.4%pD
$SFUD_{3B}$	D_{pCT}	4.9±1.1%pD	9±2%pD	2.7±0.8%pD	12±3%pD
	D_{CBCT}	4.3±1.1%pD	6.1±0.5%pD	1.4±0.5%pD	8.3±0.9%pD
	D_{dCT}	2.8±0.6%pD	3.4±0.2%pD	1.0±0.3%pD	7.1±0.5%pD
$IMPT_{5B}$	D_{pCT}	5.0±1.2%pD	9±3%pD	7±3%pD	11±4%pD
	D_{CBCT}	4.4±0.5%pD	5.8±0.8%pD	3.2±1.1%pD	8.6±1.0%pD
	D_{dCT}	2.9±0.3%pD	3.5±0.2%pD	2.5±0.5%pD	7.0±0.1%pD

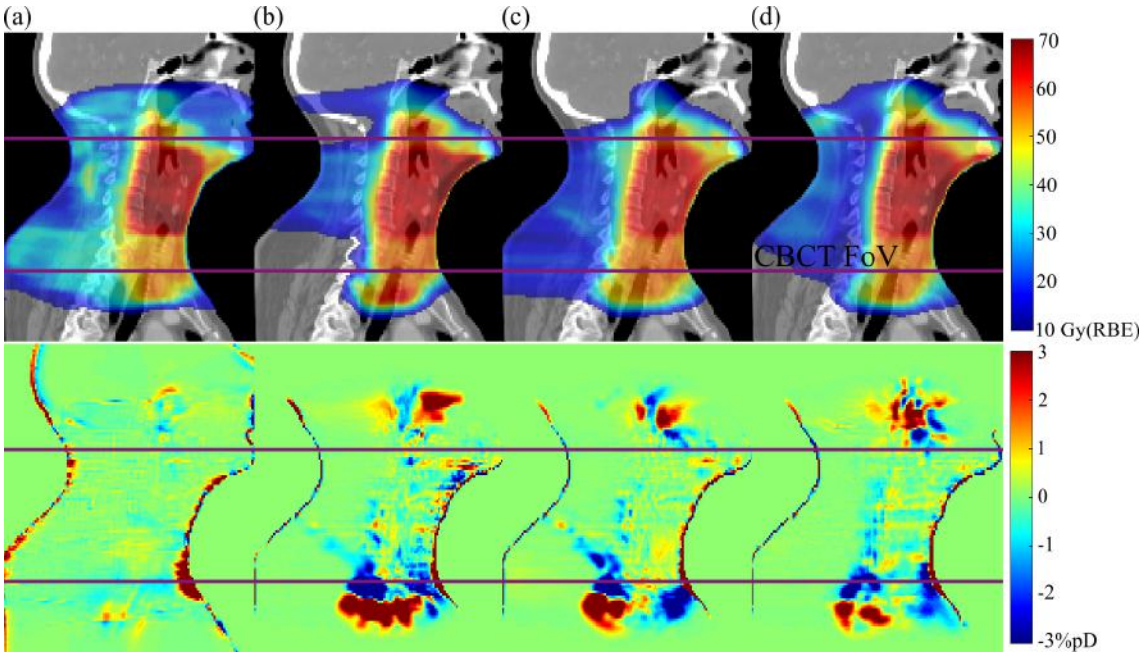

Figure 5.4: *Dose colourwash overlayed on the replan CT (top row) and difference in dose between replan CT and deformed CT (bottom row) for (a) IMRT plan, (b) $IMPT_{3B}$, (c) $SFUD_{3B}$ and (d) $IMPT_{5B}$, for one of the patients included in this study. The horizontal purple lines indicate the length of the CBCT field-of-view (FoV).*

Table 5.5: *Mean ± standard deviation of the dose difference test pass-percentage ($DD_{2\%-pp}$), mean and root mean square of the dose differences (DD_{mean} and DD_{RMS}) and differences in calculating the mean and maximum doses (ΔD_{mean} and ΔD_{max}) to OARs (spinal canal, brainstem and parotids)*

Plan	$DD_{2\%-pp}$ (%)	DD_{mean} (%pD)	DD_{RMS} (%pD)	ΔD_{mean} (%pD)	ΔD_{max} (%pD)
IMRT	99.9±0.1	0.2±0.1	0.3±0.1	0.1±0.1	0.2±0.1
$IMPT_{3B}$	81±17	1.3±1.0	2.3±1.6	0.8±0.9	1.8±1.7
$SFUD_{3B}$	85±12	0.9±0.6	1.6±0.8	0.5±0.5	1.3±2.5
$IMPT_{5B}$	80±21	1.7±2.1	3±3	1.4±2.0	1.8±1.6

SFUD optimisation and/or additional IMPT beams also lead to better performance.

Table 5.5 presents the results found for D_{dCT} within OARs. The same trend of better results being obtained for photon plans is observed here. The larger standard deviation found for $DD_{2\%-pp}$ in proton plans is due to the relative positioning of the different OAR in the dose map for each patient. For $IMPT_{3B}$ the $DD_{2\%-pp}$ was 94±6% for the brainstem and spinal canal and 68±14% for the parotids. In general, the parotids have the worst results due to being partially within the TV, and therefore being more susceptible to high doses and high dose gradients. One of advantages of using DIR to generate a dCT is to also generate deformed contours automatically, removing the need to delineate all the regions of interest from scratch. The differences when plotting the DVHs were small when the manual structure set was considered. If the pCT structures were used to generate deformed contours in the dCT, the differences between curves became more evident (Figure 5.5). The same trend was also found for IMRT plans. The variability in OAR contouring plays a significant role when evaluating the need to replan independently of the treatment modality, and therefore consistent delineations between time-points are important.

5.5 Discussion

CBCT imaging is a very common imaging modality in photon therapy, but it is still in its infancy in proton therapy. With CBCT becoming clinically available, it becomes important to understand how to use it for treatment adaptation and what the inherent uncertainties associated are. Very recent studies by Landry et al. also evaluated DIR and CBCT based proton dose calculations on a deformable phantom and patient data, but unlike this work these studies were less focused on the dosimetric implications and the impact of different treatment strategies [11, 12]. Calculating the "dose of the day" directly on the CBCT was extensively studied for photon therapy [9, 55]. Both approaches can benefit from further refining due to the greater accuracy required for proton therapy applications. Using a dCT was more accurate than a calibrated CBCT approach. Visual inspection of the doses obtained showed that larger imaging volumes have higher noise

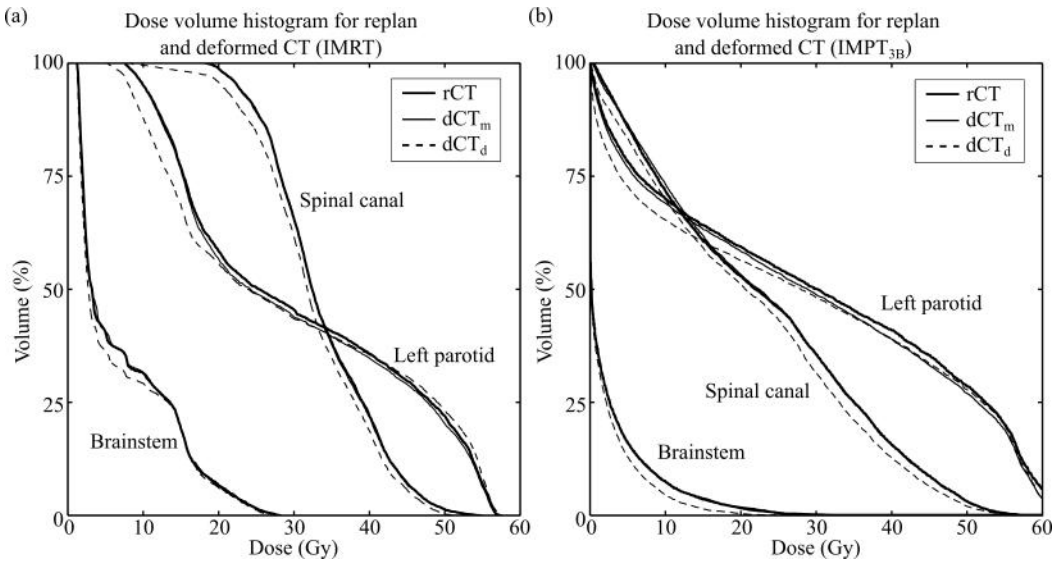

Figure 5.5: *Dose volume histogram comparing dose in replan CT (rCT) and deformed CT (dCT) for (a) IMRT and (b) $IMPT_{3B}$, using manual (dCT_m) and deformed (dCT_d) structures. Right parotid omitted for figure clarity.*

and lower contrast, and are the main source of errors in recalculating proton dose directly on a CBCT. The CBCT imaging and calibration protocol was equal to that of chapter 3 (i.e., indicative but clearly suboptimal).

The results reported here for photon plans were comparable but not equal to those reported in chapter 3 (i.e., Table 3.3). The improvement results from the higher similarity between dCT and rCT in this study, as using LNCC over NMI as similarity measure proved to result in better alignment after registration; however, there are two other sources that cause differences between the results. The cohort used here was smaller than the one of chapter 3 and the IMRT plans used were different.

The results found for the dose similarity between dCT and rCT are promising for proton therapy applications, even though significantly inferior to the IMRT cases. The estimation of the $IMPT_{3B}$ dose was less accurate in the region of high dose gradient but still showed good accuracy in the regions of highest clinical importance, (i.e., TV and OARs). The lack of a perfect gold-standard is the major limitation of studies evaluating the "dose of the day", as is the focus here and was in chapter 3. However, as there would still be setup errors between the CBCT and rCT scan even if they were acquired on the same day (as the scanners are in different rooms), the extra and unnecessary (from a clinical perspective) imaging dose to the patient needed to acquire an extra CBCT on the same day as the rCT was not justifiable, particularly for patients that were treated with IMRT originally. The effect of using an imperfect gold-standard is more likely to have a negative than a positive impact on our results, i.e., any mismatching between the rCT and dCBCT should not help the proposed method to appear better than it really is. It is true that this could happen by chance, if the errors in registering the pCT to the dCBCT are the exact opposite of the errors when generating the dCBCT, but this is very

unlikely. Furthermore, the impact of using an imperfect gold-standard was estimated by also performing the same dose analysis on a dCT resulting from warping the pCT to match the rCT (dCT_{rCT}). In such case, the geometric information is correct thus the additional uncertainty caused by the incorrect gold-standard should lay between the values reported in this study and the results found on the dCT_{rCT}. For proton plans this additional uncertainty within the CBCT FoV was estimated to be approximately 5% and 1.5%pD for $DD_{2\%-pp}$ and DD_{RMS}, respectively. An additional limitation is the small patient sample used. While it is appropriate for a proof-of-concept study such as this, follow-up studies with larger samples are required to fully characterise the uncertainty of CBCT and DIR based "dose of the day" calculations.

The method proposed is of potential interest for adaptive proton therapy, but further work is necessary to address its poorer reliability in regions where the dose varies rapidly. Three major points that are crucial for CBCT and DIR-based adaptive proton therapy workflow, and that can be further improved were identified: (1) registration algorithm, (2) CBCT acquisition for ART applications and (3) treatment plan robustness. (1) Proton plans are more sensitive to inaccuracies in the registrations. Even though the state-of-art registration algorithm available in NiftyReg was used, this is a general purpose algorithm designed to be applicable to a wide range of medical images from different modalities and of different parts of the anatomy. General purpose algorithms can be made more realistic by incorporating additional constraints (e.g., to avoid bone deformation [110]). Alternatively, algorithms specifically designed for the treatment site could be used, such as biomechanical-based algorithms that model the physical properties of the tissues being registered [160, 161]. Further work is also necessary to ensure the registration of (and attenuation by) the immobilisation devices. For simplicity reasons the immobilisation mask and treatment couch were removed from the dose calculations, but the mask can contribute significantly to the attenuation of proton beams. Furthermore, Landry et al. point also discussed the importance of properly dealing with errors in the registration of the trachea and airways, which are very common in the HN site due to swallowing motion, and were shown to locally impact the dose recalculation [12]. (2) The information acquired by the CBCT has to be adequate as missing important geometrical information due to limited FoV is concerning particularly for adaptive proton therapy. Since the length of the CBCT scan is limited by the geometry of the system, acquiring two consecutive images may be the most appropriate solution for larger treatment volumes. Further improvement of the image quality will also facilitate the registrations. (3) $SFUD_{3B}$ and $IMPT_{5B}$ dose distributions were less sensitive to registration errors than $IMPT_{3B}$, which can be related with the robustness of the plan. Robust treatment plan optimisation can result in plans that maintain target coverage and normal tissue sparing in the presence of setup errors and range uncertainties [162]. Robust planning will play an important role in minimising and accounting for the issues generated by the anatomical changes, and could be used to make plans that are less sensitive to registration errors. The level of accuracy necessary for the dose calculations will depend on the final application. For example,

if the aim is to use the "dose of the day" to identify which patients may benefit from replanning the accuracy requirements are not necessarily high as the decision to replan is being made clinically on a patient-by-patient basis. It is crucial for clinical translation to fully characterise the errors and uncertainties inherent to the dose calculations, and to develop planning methods that account for and are robust to them.

This study assumes that CBCT imaging quality is the same as if the system was mounted on a LINAC head, which may not be the case. The larger source to detector distance and the flex of the support system induced by gravity increase the magnitude of the geometric deformation in comparison with photon systems, and degrade the reconstruction process of the CBCT image [163]. The effect of such degradation will have to be verified once patient CBCT data during proton therapy becomes available. It is expected that the lower image quality will have more of an effect on dose calculations directly on CBCT than on a dCT.

Even though the method proposed is still not optimal and can certainly be improved, it is clear the importance of treatment adaptation in HN proton therapy, as dose differences between pCT and rCT show. It becomes evident that the range of proton beams within the patient needs to be predicted as accurately as possible not only during treatment planning but also throughout the treatment course. Knowledge of where the proton dose is being delivered throughout the treatment and online treatment adaptation is made possible by the introduction of CBCT imaging, which has the potential to bring additional confidence to reduce the larger safety margins characteristic of proton therapy, and therefore fully utilise the potential advantages of this treatment modality [164]. The framework here presented is not exclusive to HN and could be further extended to other anatomical sites of interest. However, site-specific validation work will be necessary as it is crucial for the DIR algorithm to be tailored for the specificities of the region being registered and the quality of the images acquired in such volumes. Chapter 6 will investigate the use of CBCT and DIR for adaptive lung proton therapy using CBCT data acquired in a proton-gantry.

5.6 Conclusions

In this chapter a feasibility study was conducted to investigate the use of CBCT and DIR for adaptive HN proton therapy dose recalculation. This was an extension to proton therapy of the work presented in chapter 3 using a more sophisticated DIR algorithm. Promising results were found even though the proposed method performed worse for proton than for photon treatments. This work allowed to identify the major areas where further work is necessary to facilitate CBCT and DIR driven ART, which include improvements to registration, image acquisition and robust planning strategies.

Chapter 6

Lung adaptive proton therapy

You may delay, but time will not.

Benjamin Franklin

The use of CBCT and DIR was investigated for dose calculations for HN in the context of IMRT and proton therapy (chapters 3 and 5). Following the installation of the world's first CBCT system at Roberts Proton Therapy Center (Philadelphia, USA), real CBCT for proton patients became available. Therefore, the natural following step of this thesis was to investigate and further develop the dCT method for adaptive lung proton therapy.

The work in this chapter resulted in the following outputs:

- *(In preparation)* C. Veiga, G. Janssens, T. Baudier, L. Hotoiu, S. Brousmiche, J. R. McClelland, C.-L. Teng, L. Yin, G. Royle, and B.-K. K. Teo, "The accuracy of CBCT and deformable registration for adaptive lung proton therapy" (2016).

- C. Veiga, G. Janssens, C.-L. Teng, T. Baudier, L. Hotoiu, Lingshu Yin, J. R. McClelland, G. Royle, C.B. Simone II, and B.-K. K. Teo, "Quantitative assessment of proton range deviations using lung CBCT," Proceedings of the 55th Annual Conference Particle Therapy Co-Operative Group (Prague, Czech Republic) (2016).

- C. Veiga, G. Janssens, C.-L. Teng, T. Baudier, L. Hotoiu, J. R. McClelland, G. Royle, L. Lin, L. Yin, J. Metz, T. D. Solberg, Z. Tochner, C. B. Simone II, J. McDonough, and B.-K. K. Teo, "First clinical investigation of CBCT and deformable registration for adaptive proton therapy of lung cancer," Int. J. Radiat. Oncol. Biol. Phys. 95(1) 549-559 (2016).

6.1 Rationale

In the context of lung malignancies, proton therapy offers better dose localisation than that achieved by conventional photon therapy [73, 165–168]. The dosimetric advantage of protons over photons has been used to improve the poor outcome of lung cancer patients [169], allowing the reduction of dose to critical structures such as the lung, heart and spinal cord [170]. PSPT has predominantly been used in the clinic as a robust strategy to deal with the intra-fractional motion aspect of lung tumours; however, other inter-fractional changes during the course of radiotherapy may also affect the dose delivered to target and healthy tissues [85, 171]. These factors include changes in tumour size and position, alterations in tissue anatomy, variations in respiratory patterns, and fluctuations in patient weight [86]. In conventional photon radiotherapy tumour regression was found to occur in 40% of patients undergoing definitive treatment [172], with reductions of up 70% of their volume reported in the literature [173]. Lung changes during the course of treatment, such as increase or decrease of atelectasis (i.e., lung collapse) and pleural effusion (i.e., liquid accumulated between lungs and the ribs or diaphragm), are less frequent but can dramatically modify the range of the proton beam. Additionally, movement of the tumour can evolve throughout the treatment fractions; therefore, re-evaluation of the targets may be required [174]. At HUP replanning is triggered in about 10-20% of the lung cancer patients treated with PSPT.

Proton dose distributions are highly sensitive to changes in patient geometry and tumour volume, especially in the lungs [175]. To fully utilise the advantage of proton therapy, positioning uncertainties and anatomical changes have to be monitored and accounted for. The position of the SOBP is sensitive to changes in tissue density along the beam path, which may result in potential shifts of the SOBP. For example, interfractional tumour enlargement or development of atelectasis increase density along the beam path and shorten beam penetration. The under-ranging can potentially reduce target coverage. Conversely, tumour regression reduces density along the beam path and increases beam penetration. The over-ranging may result in unplanned dose to otherwise spared organs distal to the tumour volume. Inter-fractional adaptive replanning is therefore beneficial to select patients [176]. Therefore, accurate patient positioning and regular evaluation CTs scans are critical components of proton therapy [170]. Replanning will be required if the new dose distribution based on evaluation CTs compromises target coverage and/or exceeds tissue tolerance. Volumetric imaging afforded by CBCT is an alternative to routine CT imaging and, as previously seen for HN patients, may play an important role in lung adaptive proton therapy.

Gantry mounted CBCT systems are now available for image-guidance in proton therapy. The world's first CBCT mounted in a proton-gantry was installed in HUP, and the first patient scanned in August 2014. dCT was shown in chapter 5 and other literatures studies [11, 12] to be a potential surrogate for repeat CTs for proton treatment verification

in the context of HN malignancies; however, these retrospective studies used data from LINAC CBCT systems.

This chapter is focused on two complementary studies. First, a comprehensive evaluation of the uncertainties associated with an adaptive proton therapy workflow based on CBCT, DIR and a fast range-corrected estimate of the "dose of the day" was performed. The accuracy of the dCT was assessed in terms of water equivalent thickness (WET) and "dose of the day" estimation. Secondly, a clinical ART workflow using on-board CBCT where a replan is triggered after three decision points was proposed. This workflow was evaluated in terms of clinical indicators of replanning for a diverse cohort of lung cancer patients summarising common radiation-induced changes in the lung. A rCT was used as the gold-standard to gauge the accuracy of the dCT for both studies. This was both the first study focused on quantifying the accuracy of DIR and CBCT for ART, and evaluating the proposed clinical workflow in lung proton therapy.

6.2 Methods and Materials

6.2.1 Patient selection and data acquisition

Data from twenty consecutive patients treated for lung malignancies were included in this retrospective study. All patients underwent PSPT using two treatment fields with a median dose of 66.3 Gy(RBE) (range: 40-66.6 Gy(RBE)) in a median of 1.8 Gy(RBE) per fraction (range: 1.5-4 Gy(RBE) per fraction). The patient cohort included a variety of tumour sizes, locations and anatomical changes that occurred throughout the treatment course (Table 6.1). Those changes included dramatic changes in the lung (10%), such as atelectasis and lung reinflation, small changes in tumour and setup variations (35%), moderate (<25% of the GTV volume) tumour shrinkage/enlargement and drift in tumour location (35%) and large (>25% of the GTV volume) tumour shrinkage (20%).

The imaging protocol consisted of a 4D PET/CT (Gemini TF Big Bore PET/CT, Philips Healthcare, Andover, MA, USA) for treatment planning, proton-gantry mounted CBCT (Ion Beam Applications SA, Ottignies-Louvain-la-Neuve, BE) and rescan 4D CT (Sensation Open, Siemens Healthcare, Malvern, PA, USA) acquired in treatment position for plan verification during the course of treatment. The average of the 4D CT was used for treatment planning and to generate the CTs. One pair of CBCT and rCT at mid-treatment were selected for evaluation for each patient. The chosen rCT and CBCT scans were acquired with up to 2 days difference (85% same day, 5% 1 day difference, 10% 2 days difference) [1]. The CBCTs were acquired using a half scan mode at 110 kVp and 1142 mAs. The CBCT system has a source-to-axis distance of 288.4 cm, detector-to-axis distance of

[1] Unlike with the datasets of the HN region used in this thesis, at HUP it was approved to acquire CT and CBCT scans at the same day. This additional imaging was allowed as part of the commissioning of the first proton-gantry mounted CBCT system.

Table 6.1: *Characteristics of the patients included in this study: age, TNM staging, treatment fields, target volumes, fraction at which the CBCT was acquired, and anatomical changes detected at the verification scan.*

PT #	Age (Y)	G	TNM	Fields (gantry angle/name)	V_{iCTV}[a]	F_{rCT}/F_t[b]	Characteristics at verification scan	Tumour motion: SI/LR/AP[c] (mm)	
								pCT	rCT
1	85	M	T3N2M0	175(LPO_1) & 135(LPO_2)	130	9/37	Atelectasis	6/6/4	6/7/5
2	71	F	T3N2M0	210(RPO_1) & 220(RPO_2)	410	20/37	Large tumour shrinkage[d]	8/5/5	10/4/4
3	68	F	T4N2M0	220(RPO) & 180(PA)	280	4/30	Lung reinflation	10/4/10	12/3/2
4	69	M	T4N2M0	180(PA) & 220(RPO)	280	21/37	Small tumour shrinkage/setup error	5/1/3	5/1/3
5	72	M	T3N3M0	180(PA) & 10(RAO)	260	8/27	Tumour position drift	9/2/4	8/3/2
6	81	M	T4N0M0	180(PA) & 155(LPO)	320	9/37	Small tumour shrinkage/setup error	3/1/2	2/1/1
7	77	M	T2aN2M0	180(PA) & 200(RPO)	180	10/37	Small tumour shrinkage/setup error	7/3/2	5/2/2
8	62	F	T4N2M0	270(ASO) & 180(PA)	340	4/34	Large regression of infiltrating tumour	2/3/6	4/2/5
9	64	M	T2aN0M0	180(PA) & 155(LPO)	130	22/37	Small tumour shrinkage/setup error	4/2/2	4/1/3
10	65	M	T4N1M0	180(PA) & 150(LPO)	180	10/25	Small tumour shrinkage/setup error	5/3/4	8/3/4
11	31	M	T2aN2M0	0(AP) & 205(RPO)	200	20/37	Moderate tumour shrinkage[e]	7/2/5	7/4/4
12	76	F	T2bN0M0	180(PA) & 210(RPO)	150	20/37	Moderate tumour shrinkage	2/0/0	2/0/1
13	71	M	T3N2M0	180(PA) & 145(LPO)	190	21/25	Moderate tumour shrinkage	6/3/3	6/3/6
14	65	F	T4N0M0	180(PA) & 150(LPO)	500	20/37	Large regression of infiltrating tumour	2/2/2	1/0/0
15	57	F	T3N2M0	180(PA) & 225(RPO)	340	10/30	Moderate tumour shrinkage	3/2/0	1/0/0
16	58	F	T4N0M0	205(RPO) & 20(LAO)	580	15/33	Moderate tumour enlargement	6/6/4	2/1/1
17	62	F	T1aN2M1b	175(LPO_1) & 150(LPO_2)	100	5/37	Moderate tumour shrinkage	2/0/2	3/1/4

18	67	M	T3N1M0	0(AP) & 335(RAO)	330	20/30	Small tumour shrinkage/setup error	3/3/3	5/3/3
19	69	M	T3N0M1b	180(PA) & 205(RPO)	310	7/16	Small tumour shrinkage/setup error	2/2/2	1/1/1
20	78	M	T2aN0M1b	195(RPO_1) & 210(RPO_2)	140	9/15	Large tumour shrinkage; density changes	5/2/2	5/4/3

[a] V_{iCTV}=volume of the iCTV

[b] F_{rCT}/F_t=treatment fraction at which rCT was acquired/total number of fractions

[c] SI=superior-inferior; LR=left-right; AP=anterior-posterior

[d] Large tumour regression = visually apparent change in tumour volume greater than 25% of its original GTV

[e] Moderate tumour regression = visually apparent change in tumour volume less than 25% of its original GTV

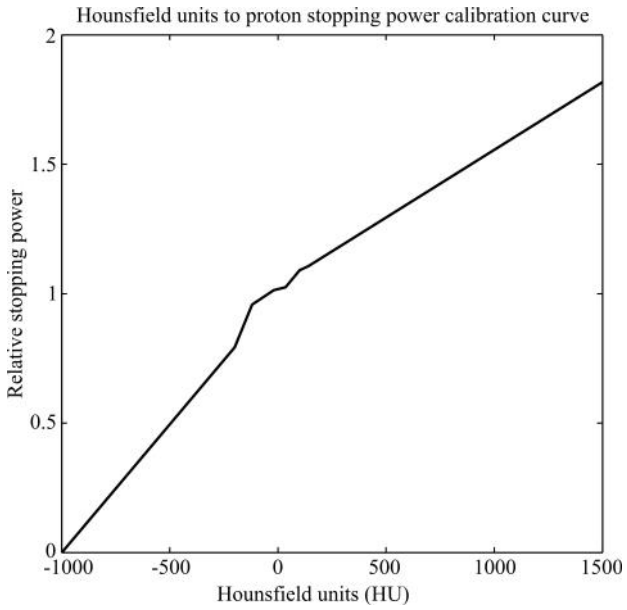

Figure 6.1: *Relative stopping power calibration curve imported in the clinical treatment planning system.*

58.6 cm and a maximum FoV of 34 cm in diameter and length. The image resolutions were 1.17×1.17×3.0 mm^3, 1.33×1.33×2.5 mm^3 and 0.98×0.98×3 mm^3 for the pCT, CBCT and rCT, respectively. The open-source Reconstruction Toolbox (RTK) was used for CBCT reconstruction [177]. The two CT scanners had the same stopping power calibration curves for proton dose calculations (Figure 6.1).

6.2.2 Overview of an adaptive lung proton therapy workflow

A clinical ART workflow using on-board CBCT is proposed, where the replan process is triggered after three decision points (Figure 6.2). The first is a fast online process, where a range-corrected dose distribution based on WET is rapidly calculated on a dCT derived from the CBCT. When significant dosimetric changes are observed, an offline review process is triggered for a full dose recalculation on the dCT. If the dosimetric impact is still evaluated as significant, a rCT is scheduled. If dosimetric changes are confirmed on the rCT, a replan is triggered. The online decision point is based the following steps:

1. The pCT is deformed onto the CBCT acquired for treatment verification, creating a dCT. This dCT contained the HU information of the pCT and the geometry of the daily anatomy.

2. DIR alone cannot account for all the complex changes that occur in the region of the thorax, which include changes to both lung and tumour. When appropriate, the dCT was corrected for gross registration failures based on the difference between dCT and CBCT.

3. The WET was calculated in the anatomy along the beam path for each field and used to rapidly estimate the range and dose distribution of the treatment fields.

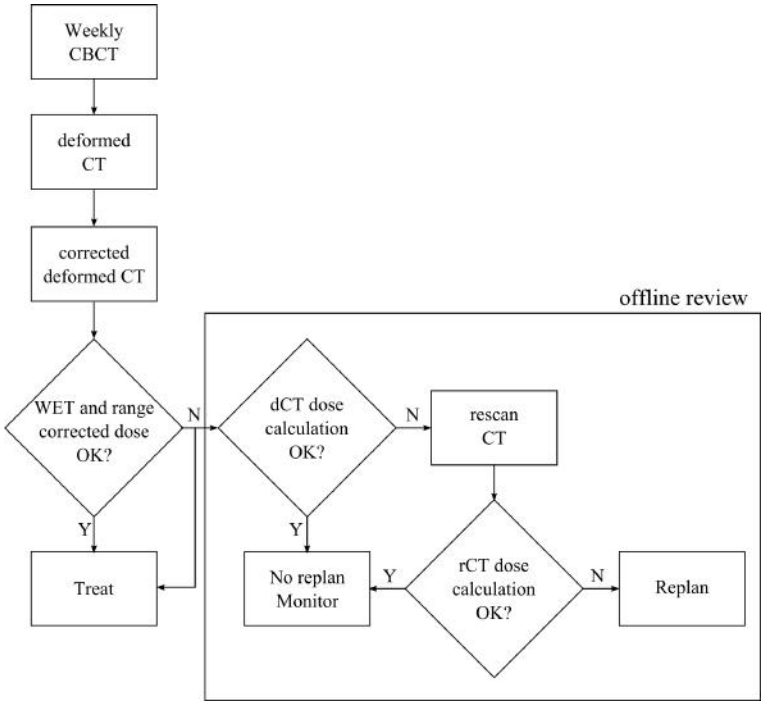

Figure 6.2: *Workflow for ART. A corrected dCT was created using pCT-to-CBCT DIR, and the variation in WET between dCT and pCT was used to range-correct the planned dose. This process can be performed online to trigger an in-depth offline review of the dCT (and rCT) if deemed necessary.*

Variation in WET on the distal surface of the target is also a good surrogate for potential under/over-ranges.

4. Clinical indicators based on under/over-ranges, underdosage of target volumes and/or overdose of OARs are then calculated to trigger the offline workflows (i.e., investigate the need for a rCT and subsequent replanning).

6.2.2.1 Deformable registration

The first step of the workflow is to register the pCT to the CBCT. A rigid registration was first applied in order to estimate the global alignment between the pCT and the CBCT. The rigid registration was a manual alignment in the treatment planning system of the images based on the same criteria used during treatment for correction of the patient positioning, i.e., matching bony anatomy at the target volume level. The obtained transformation was then used to initialise the DIR. The couch was delineated and removed from the images prior to DIR in order to avoid any influence on the registrations. The DIR was implemented using the Morphons algorithm available in REGGUI [178]. Because this algorithm is based on the matching of local phases, it was robust to changes in intensity when applied to CT images of the thorax with and without contrast enhancement [41]. In the present study, the Morphons algorithm used eight resolution scales with ten iterations for the six coarsest scales, five and two for the second finest and finest scales, respectively. A Gaussian regularisation of 1.25 mm standard deviation was used.

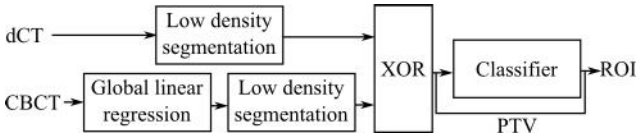

Figure 6.3: *General pipeline for the detection of regions of interest for dCT correction. The pipeline is initialised by correcting the intensities of the CBCT via a global linear regression. The low density regions are then segmented on both images using a watershed-cuts algorithm. An exclusive OR logical operator then compares the segments to identify areas where tissue density has changed. A Perceptron based classifier compares the detected areas and classifies regions of interest (ROIs). Detected areas within the PTV are automatically classified as ROI.*

6.2.2.2 Deformed CT correction

Some anatomical changes in the thorax cannot be modelled by deformation alone. DIR algorithms use a transformation model that represents expansion and contraction of tissues, and therefore are not adequate in situations where tissue appears or disappears. The clinical situations are diverse and include changes within the lung (such as atelectasis and pleural effusion) and complex tumour change in response to treatment (such as regression of infiltrating tumours and erosion [179]). A pipeline that identifies problematic regions needing further corrections for gross registration failures was developed (Figure 6.3). This method detects low density differences between the dCT and CBCT as they are the type of issues attempted to be corrected. The pipeline is initialised by correcting the intensities of the CBCT via a global linear regression between CBCT and dCT intensities. The low density regions are then segmented on both dCT and CBCT using a watershed-cuts algorithm [180, 181]. An exclusive OR logical operator then compares the segments to identify areas where tissue density has changed between dCT and CBCT (e.g. atelectasis). A Perceptron based classifier [182, 183] compares the detected areas and classifies region of interests (ROIs) with significant mismatch between dCT and CBCT. The classifier uses the decimal logarithm of the ROI size and the mean absolute difference between CBCT and dCT within the ROI to define the acceptability limit. Detected areas within the PTV are automatically classified as ROI because they have an important impact on the dose calculations. In this current prototype version of the workflow, the ROI step was implemented in a semi-automatic fashion, i.e., the regions of mismatch are automatically identified but the user has the opportunity to discard any ROIs that were incorrectly detected. In general, gross errors are easy and quick to spot, as most of the false detections occur outside the lung due to scatter artefacts. The user can also dynamically relax the acceptability criteria thus allowing more regions to be detected. Furthermore, the lung contours and target volumes can be used to automatically discard ROIs outside the lung. Within the ROI, the dCT pixel values were replaced by the bulk value of lung or tissue based on thresholding of the intensity-corrected CBCT, i.e. if the CBCT intensity is lower than -400, the dCT was replaced with value of lung (-650 HU); otherwise, it was replaced with value of water (0 HU). These values were empirically chosen based on the average intensity of lung and tissue on the CBCTs and CTs.

6.2.2.3 Water equivalent thickness

WET represents the equivalent thickness of water that would cause a proton beam to lose the same amount of energy in a given thickness of a different medium. The WET was calculated per beam angle on both the real and deformed CT scans using the following steps. The CT HUs were first converted to relative stopping power using the tissue-substitute calibration method [184]. The WET map was then obtained by accumulating the relative stopping power on a voxel-by-voxel basis of all tissues crossed by an infinitely thin beam reaching the point from a virtual proton source. In the current implementation, the WET was calculated for every voxel in the body (i.e. 3D WET map), as well as projected into surfaces with respect to the beam's eye-view (i.e. 2D WET maps).

6.2.2.4 Range-corrected dose

For a fast online workflow to estimate the "dose of the day", a range-corrected dose approximation method similar to the one described by Park et al. was implemented [185]. The method relies on the assumption that anatomical changes impact only the in-depth distribution of the dose and not its lateral distribution. The "dose of the day" is approximated by warping the planned dose based on the in-depth mapping between the anatomy at planning and the anatomy of the day. The mapping is derived from the WET maps computed from the pCT and the dCT. As a result of the warping, the dCT dose in one voxel of given WET in the dCT-based WET map is equal to the pCT dose at the position along the beam path of equal WET in the pCT-based WET map. An additional correction is applied based on the source-to-axis distance to account for the loss in protons in a divergent beam. This method estimates the impact of the anatomical change in the planning dose distribution without requiring a traditional dose recalculation on the treatment planning system. This algorithm was benchmarked against a dose recalculation using the clinically commissioned proton therapy treatment planning system at HUP (Eclipse, version 11.0, Varian Medical Systems, Palo Alto, CA, USA). A sample of 40 dose distributions pairs (range-corrected and recalculated) were used in this benchmarking.

6.2.2.5 Clinical indicators

The WET and dose 3D information can be used to automatically generate a set of clinical indicators to estimate the impact of anatomical changes on the treatment objectives, and thus aid the decision-making process.

Variation in WET on the distal surface of the target is a good surrogate for potential under/over-ranges. Changes in range were estimated by computing the difference between the WET from the pCT and dCT (i.e. $WET_{pCT} > WET_{dCT}$ corresponds to over-ranging and $WET_{pCT} < WET_{dCT}$ to under-ranging). However, in the case of PSPT, compensator smearing is used to ensure target coverage in presence of errors in patient

positioning and motion [186]. Therefore, potential under-ranges are partially taken into account by the compensator (which increases the beam range); thus, for under-ranges a morphological dilation using the same radius as the compensator is applied on the pCT-based WET map before computing the difference with the dCT-based WET map to identify under-ranges not accounted for in smearing. Quantitative measurements of the 2D WET difference maps at the PTV distal surface (per beam) for the pCT and dCT were calculated as clinical indicators, which include the percentage of pixels with under/over-ranges larger than 3 mm ($WET_{under>3mm}$/$WET_{over>3mm}$) and the 95% percentile of the under/over-range distribution ($WET_{under\text{-}95\%}$/$WET_{over\text{-}95\%}$).

DVHs and dosimetric statistics representative of target coverage and OAR tolerances congruent with HUP guidelines to trigger replanning after offline review of lung patients were automatically generated by the workflow. These were for PTV and iCTV the $V_{95\%}$ and $V_{99\%}$, respectively, with a threshold of 3% change in the rCT. Similarly, for OARs the dose tolerances used were: for heart, D_{max}=72 Gy, V_{45Gy} <35% and V_{30Gy} <50%; for oesophagus, D_{max}=70 Gy and V_{55Gy} <30%; for cord (and cord+5mm), D_{max}=50 Gy(65 Gy); and for brachialplexus, D_{max}=66 Gy. The iCTV/PTV contours were rigidly propagated, even when the tumour regresses. The OAR contours were propagated automatically to the verification scans using DIR.

6.2.3 Implementation details

The REGGUI package was used to evaluate the proposed adaptive radiotherapy workflow. REGGUI is a MATLAB (MathWorks, Natick, MA, USA) based image processing software featuring various registration methods, filtering methods, segmentation tools and other radiotherapy dedicated functions such as CBCT simulation, DVH computation and gamma index computation. The software was designed in a way that facilitates the construction of clinical data workflows for research purposes. Several of the different algorithms described previously were implemented in REGGUI particularly for this project. The WET calculation and range-corrected dose calculation were implemented by Dr Guillaume Janssens (IBA); the dCT correction step was implemented by Thomas Baudier and Dr Lucian Hoitou (iMagX Project, Université Catholique de Louvain); and the CBCT simulation process was implemented by Dr Guillaume Janssens and Dr Sebastien Brousmiche (IBA). I contributed closely in the design, optimisation and validation of those tools, and was mainly responsible for the implementation of the pipelines in REGGUI; furthermore, I also developed a number of auxiliary functions necessary for complete integration of the workflows. The pipeline of processes implemented in the adaptive proton therapy workflow are schematically presented in Figure 6.4.

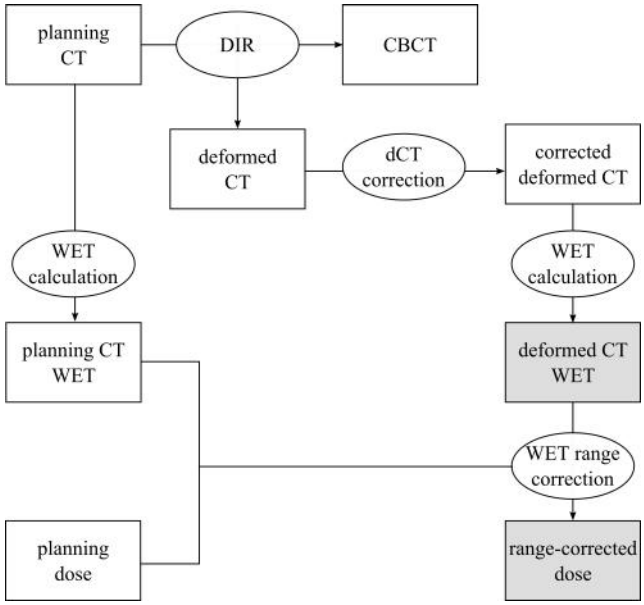

Figure 6.4: *Diagram of data and processes implemented in REGGUI for the online decision-point of an adaptive lung proton therapy workflow. The uncertainty in WET and range-corrected dose calculations (boxes highlighted in grey) were assessed by comparison with those computed on the rCT.*

6.2.4 Evaluation of the adaptive proton therapy workflow

In the context of adaptive lung proton therapy, two complementary studies were performed to assess the workflow proposed:

1. Accuracy of CBCT and DIR for adaptive lung therapy: this consisted of comprehensive evaluation of the uncertainties in WET and estimation of the "dose of the day" associated with the online decision-point of the adaptive proton therapy workflow proposed (i.e., uncertainties associated with the use of a corrected dCT).

2. Clinical investigation of replanning indicators: the workflow was evaluated in terms of the clinical indicators of replanning extracted from the corrected dCT for a diverse cohort of lung cancer patients summarising common radiation-induced changes.

The details of each study are described in the following sections.

6.2.5 Accuracy of cone-beam CT and deformable image registration for adaptive lung therapy

6.2.5.1 Deformable registration

The ART workflow implemented for clinical use at HUP utilised the DIR algorithm implemented in REGGUI. However, to evaluate the reproducibility of the method proposed a total of three DIR softwares were used in this study: the open-source softwares REGGUI

and NiftyReg, and the commercial hybrid deformable registration tool in RayStation (version 4.5, RaySearch Laboratories, Stockholm, SE). The settings used for each algorithm were optimised empirically for the thoracic region.

The parameters used for REGGUI were previously described in section 6.2.2.1. The asymmetric DIR algorithm available in NiftyReg was used with NMI as similarity measure, three resolution levels, a CP grid size of 15×15×9 mm^3, and two regularisation terms (the BE and the JL, to facilitate smooth deformation fields and penalise large volume changes, respectively). Unlike chapter 5, NMI was preferred over LNCC due to the larger variability in HU for the CBCT available from the proton system. The choice of algorithm and similarity measure allowed the use of the GPU implementation, as the workflow was aimed to be used rapidly online. Both REGGUI and NiftyReg were previously evaluated in the context of HN adaptive proton therapy (chapter 5 and other publications [11, 12]). The hybrid free-form registration algorithm in RayStation 4.5 used an intensity based similarity measure, two resolution levels of 5 and 3 mm voxel size. The maximum number of iterations was set to 1000 per resolution scale.

6.2.5.2 Cone-beam CT dataset definition

In order to demonstrate that the dCT can provide similar clinical information as the rCT, it is necessary to quantify the uncertainties associated with DIR and WET-based range-corrected dose. Ideally, the rCT would be the gold-standard to assess the accuracy of the dCT generated from the CBCT. However, the CBCT and rCT were not acquired simultaneously and would display differences due to setup as well as respiratory variations. For the purpose of decoupling these effects, two synthetic CBCT datasets were generated in addition to the regular CBCT (rCBCT): a simulated CBCT (sCBCT) and a dCBCT.

Regular CBCT rCBCTs were acquired within two days the rCT. The rCBCTs have the characteristic image quality and real limitations of the cone beam geometry, namely variability in HU and reconstruction artefacts (e.g., scatter, streaks and distortions caused by the couch). The main disadvantage was setup variations between rCT and rCBCT. The anatomical mismatch for some patients was considerable which limited the quantitative assessment of DIR accuracy.

Simulated CBCT In order to create an anatomically matched CBCT and rCT pair with no setup variations, a sCBCT was generated from each rCT. The RTK simulated raw 2D projection data using the geometry and acquisition parameters of the rCBCT for lung patients, and image reconstruction [177]. To further mimic the rCBCT image characteristics, scatter and Poisson noise were added to the raw 2D projections. A scatter kernel superposition method was used [187–189] which parameters have been estimated through

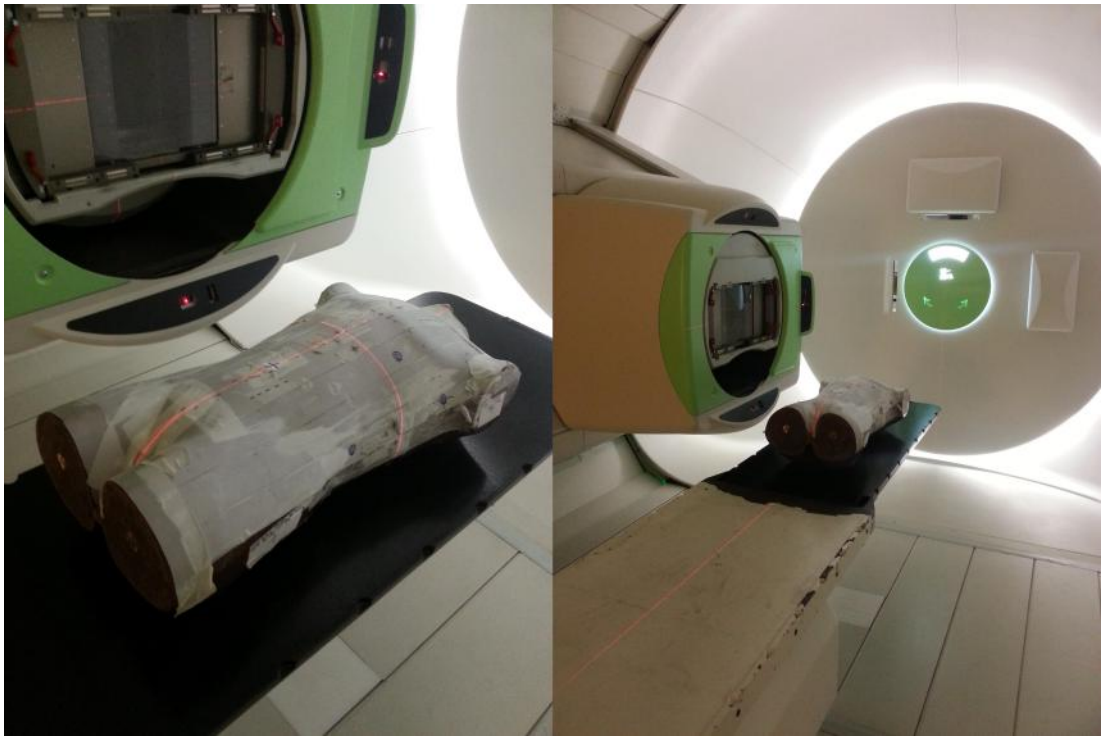

Figure 6.5: *Photos of the RANDO phantom setup for CBCT imaging.*

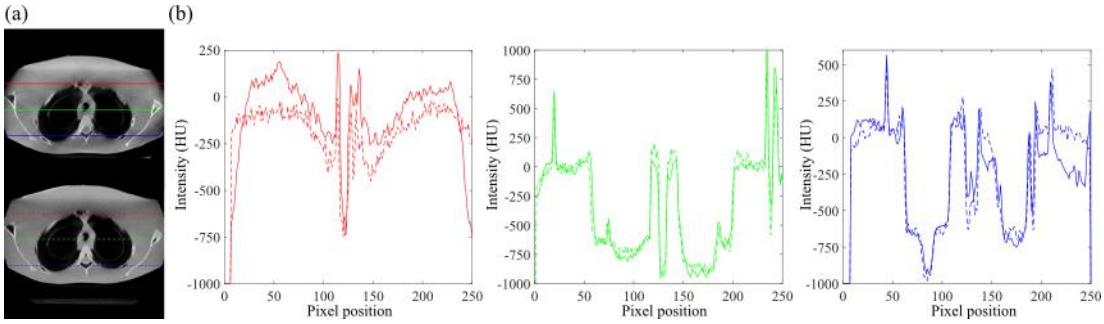

Figure 6.6: *RANDO phantom (a) rCBCT (top) and sCBCT (bottom); and (b) corresponding profiles over the horizontal direction.*

Monte Carlo simulations (GATE) [190]. The weight of these parameters was empirically fine-tuned and tested on the RANDO phantom (The Phantom Laboratory, Greenwich, NY, USA) (Figure 6.5). The RANDO phantom imaging was acquired with the collaboration of Dr Lingshu Yin (Radiation Oncology, HUP). The sCBCTs do not reproduce all the features in the rCBCT exactly. For example, scatter artefacts appeared more severe when farther away from the central region of the volume, cupping artefacts were slightly different, and streak artefacts and distortions caused by the couch were not reproduced (Figure 6.6). To investigate its impact, streaks were artificially added to the sCBCT.

Deformed CBCT Another way to create an anatomically matched CBCT and rCT pair was to use the Morphons DIR algorithm to deform the rCBCT to match the rCT. The resulting dCBCT contained the HU information of the rCBCT and the geometry of the

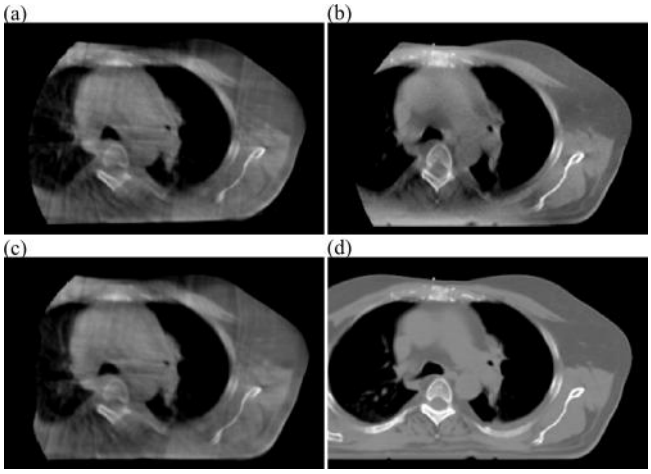

Figure 6.7: *Example of (a) rCBCT, (b) sCBCT, (c) dCBCT and (d) rCT for one of the patients included in this study.*

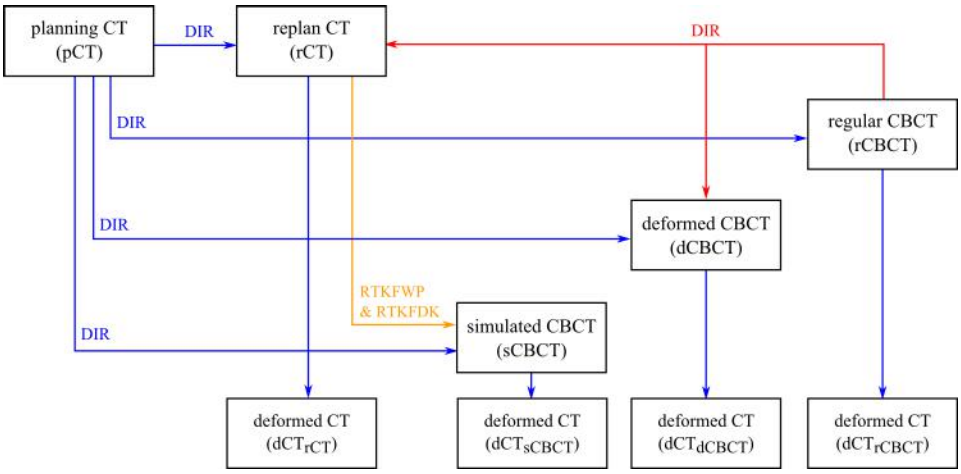

Figure 6.8: *Diagram of the different data and registrations used in this study. Three CBCT datasets were used to generate the dCTs: rCBCT, sCBCT and dCBCT. Additionally, an additional dCT was generated by deforming the pCT directly to the rCT to separate the limitations of the registration process from the limitations of CBCT imaging.*

rCT. This process introduces small geometric errors but recreates most of the artefacts present in real CBCT imaging.

The three CBCT datasets (rCBCT, sCBCT and dCBCT) were employed to generate three dCTs (dCT_{rCBCT}, dCT_{sCBCT} and dCT_{dCBCT}, respectively), which were used to report the errors associated with DIR and CBCT imaging in WET and dose estimation. Figure 6.7 provides an example of the different CBCT datasets, and illustrates the associated limitations. To distinguish between intrinsic limitations of the registration process from the inferior image quality of CBCT an additional dCT, hereby denominated as dCT_{rCT}, was generated by deforming the pCT directly to the rCT. The dCT_{rCT} was cropped to have the same FoV as the other dCTs was generated, so that only the image quality plays an impact in the validation process. Figure 6.8 summarises how the different four dCTs were produced.

6.2.5.3 Validation workflow of the deformed CT method

The accuracy of WET calculation and WET-based dose estimation in the dCT were compared for every patient, treatment field, registration algorithm and CBCT datasets.

Variation in WET on the distal surface of the target is considered a good surrogate for potential under-ranges or over-ranges. The under-ranging can potentially reduce target coverage and/or increase dose to normal tissue proximal to tumour volume. The over-ranging may result in unplanned dose to otherwise spared organs distal to the tumour volume. Therefore, the WET calculation error was calculated in both 3D (PTV) and 2D (distal and proximal surfaces), and was defined as the voxelwise difference in WET between dCT and rCT. The mean, RMS and percentile 95% of the WET differences (WET_{mean}, WET_{RMS}, and $WET_{95\%}$) were compared.

The error in range-corrected dose was defined as the difference between the obtained dose and the full dose recalculation using the TPS in the rCT; this was assessed in terms of the similarity between isodose curves, voxel-wise dose differences and a gamma-test (3%/3mm criterion). This includes the DSC for the 50% and 90% isodose volumes, and the DT between the respective surfaces (mean, RMS and 95% percentile, i.e., DT_{mean}, DT_{RMS} and $DT_{95\%}$). The mean, RMS and percentile 95% of the difference between dose distributions (DD_{mean}, DD_{RMS} and $DD_{95\%}$, respectively) were computed; for the gamma-test the test pass-percentage was calculated. Only voxels with dose larger than 20% of the maximum dose were considered. The DD results were normalised to the maximum dose (%mD). The doses were calculated and compared per field.

The treatment couch was considered a constant element (i.e., it exists in all time points in the same position), and therefore was ignored when calculating the uncertainties of the workflow.

6.2.5.4 Comparison of the deformed CT method to simpler methods

The uncertainties in WET associated with the dCT method were compared to those achieved with three simpler methods: (ii) pCT rigidly-aligned to the CBCT; (iii) bulk-density intensity assignment on the CBCT; and (iv) intensity-corrected CBCT (Figure 6.9). The main objective of this part of the study was to estimate the benefits in using the computationally expensive dCT method described previously, in comparison to other simpler methods.

Automatic tresholding directly on CBCT results in large errors due to scatter artefacts; therefore, and in place of laborious manual delineation on the CBCTs, the bulk-density intensity assignment on the CBCT was estimated using automatic thresholding of the rCT. The intensity-corrected CBCT was based on the Catphan (The Phantom Laboratory, Greenwich, NY, USA) data. The conversion between CT and CBCT was approximated by a quadratic polynomial relationship. To minimise the differences in geometric information

Figure 6.9: *Illustrative example of (a) rigidly-aligned planning CT, (b) virtual CT, (c) intensity-corrected vCBCT, (d) bulk-density assignment and (d) rescan CT.*

when comparing the different methods, only the dCBCT was used to generate the dCT and the intensity-corrected CBCT.

6.2.6 Clinical indicators of replanning

The workflow was evaluated in terms of clinical indicators of replanning. WET and dose estimation (section 6.2.2.5) were generated for the dCT and rCT, and then assessed to investigate the impact of the changes on the plan objectives. The aim was to evaluate if the dCT can be used to detect changes similarly to what would be achievable with the rCT, and thus the clinical indicators were compared to those generated from the rCT. In this study, only the dCTs generated using the rCBCT and the Morphons DIR algorithm were investigated. Hereby, when investigating the clinical indicators the general use of dCT refers specifically to $\text{dCT}_{\text{rCBCT}}$.

WET and dose estimation on the daily anatomy can be used to generate clinical indicators (section 6.2.2.5). Two dose distributions were assessed on the dCTs: range-corrected dose estimation, and full recalculation on the TPS. The range-corrected dose allows the

online portion of the workflow to be performed prior to treatment; the recalculated dose allows to carefully verify the real impact of the changes detected online. Doses were recalculated using Eclipse (version 11.0, Varian Medical Systems, Palo Alto, CA, USA). The following abbreviations are used to identify the different doses: pCT dose (D_{pCT}), dCT range-corrected dose ($D_{dCT\text{-}WET}$), dCT recalculated dose (D_{dCT}) and rCT recalculated dose (D_{rCT}). These final doses consisted of the sum of the contribution of two proton fields per plan.

The CBCT couch was, in this study, added back to the dCT after registration. The CBCT couch influences the dose distribution, and therefore is fundamental to evaluate clinical indicators of replanning (unlike to evaluate the uncertainties).

6.3 Results

6.3.1 Accuracy of cone-beam CT and deformable image registration for adaptive lung therapy

Figure 6.10 presents the DIR results for three patients in the rCT frame of reference using a color overlay with the different CBCT types. These patients represent different clinical scenarios found in the patient cohort: small tumour change and setup error (PT#6), moderate tumour shrinkage (PT#13), and non-deformable lung change (lung reinflation, PT#3).

The range-corrected dose method was benchmarked against a full dose recalculation. The average passing rate of a 3D gamma analysis (3%/3mm criteria) was 92±2%, with an uncertainty in the location of the 50% and 90% isodose surfaces of 0.7±0.5 and 0.5±0.3 mm, respectively. A qualitative comparison between planned, range-corrected and recalculated dose is shown in Figure 6.11, including over-ranges caused by tumour regression, under-ranges caused by atelectasis and little change in range. The range-corrected dose may not perfectly match the recalculated dose, but provides similar information of over and under-ranges.

6.3.1.1 Overall uncertainty of the deformed CT on water equivalent thickness and dose estimation

The overall uncertainties in WET and dose associated with the dCT were computed with the three CBCT datasets and three DIR algorithms (i.e., nine dCTs per patient). Table 6.2 presents the difference in WET calculation between the dCT and the rCT on the distal and proximal surfaces, and within the PTV. The results were averaged over all patients, fields, registration algorithms and CBCT datasets.

Table 6.3 shows the pCT dose and the range-corrected dose on the daily anatomy

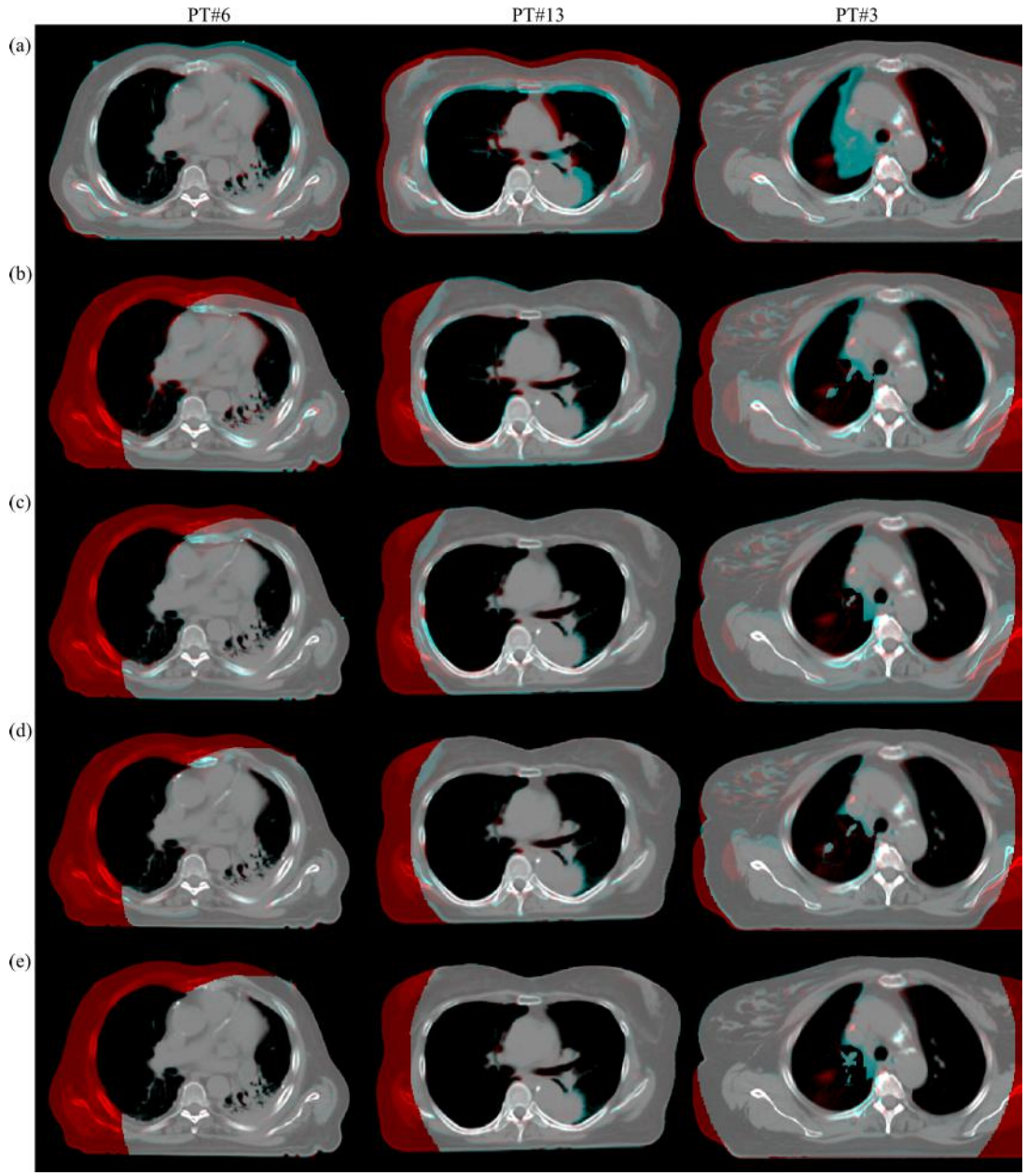

Figure 6.10: *Color overlay between rCT (red) and (a) pCT, (b) dCT_{rCBCT}, (c) dCT_{sCBCT}, (d) dCT_{dCBCT} and (e) dCT_{rCT} (cyan) for three different patients included in this study. PT#6 is a case of small shrinkage/setup errors, PT#13 of moderate tumour shrinkage and PT#3 of lung reinflation. For PT#3 the dCT correction step was necessary. For this large patient, the limited FoV cropped the external contour on the right side, which introduced registration errors of the soft tissues near the external boundaries. The dCTs were generated using REGGUI registration algorithm.*

Table 6.2: *Overall uncertainty in WET (dCT vs rCT) within the planning target volume (PTV), and on the distal and proximal surfaces (mean value ± standard deviation). The results were averaged over all patients, fields, registration algorithms and CBCT datasets.*

Region of interest	WET_{mean} (mm)	WET_{RMS} (mm)	$WET_{95\%}$ (mm)
Distal surface	0.5±2.4	3.9±2.1	8±4
Proximal surface	0.2±2.0	2.5±1.5	5±3
PTV	0.4±2.4	3.5±2.0	7±4

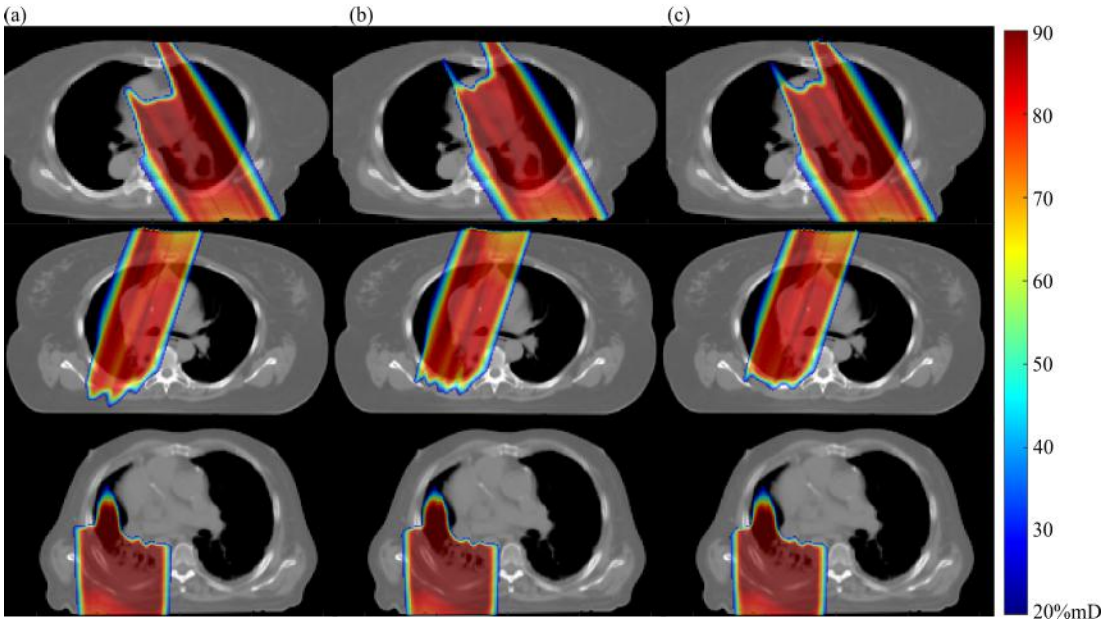

Figure 6.11: *Dose colorwash overlay on rCT using (a) pCT dose distribution from TPS, (b) range-corrected dose based on dCT_{dCBCT} WET changes, and (c) rCT dose recalculation using TPS. Examples of over-ranges (top row), under-ranges (middle row) and no change (bottom row). The dCTs were generated using REGGUI as registration algorithm.*

(dCT), in comparison with the recalculated TPS dose using the rCT. The comparison with the planned dose provides a baseline of the DDs that occur during treatment. On average, only a small difference was observed between pCT or dCT versus rCT dose. The standard deviation over the patient population is considerably higher for the pCT dose (Figure 6.12), indicating that some patients had larger differences between the pCT and rCT. In general, dosimetric parameters are less sensitive to anatomical change when multiple fields were used. Summing the contribution of each field averaged out the differences. The DD_{RMS} decreased to 4.6±1.9 and 7±4 %mD for dCT and pCT from single to two proton fields.

6.3.1.2 Effect of different cone-beam CT datasets

The unique features of the dCT generated by each CBCT dataset were used to decouple the different sources of variation in the WET and range-corrected dose uncertainty (Table 6.4). The dCT_{rCBCT} exhibited the largest differences caused by the anatomical mismatch between the rCBCT and rCT. For dCTs generated from the sCBCT and dCBCT only (i.e., ignoring the dCTs from rCBCT), the overall WET_{RMS} measured in the distal surface decreased from 3.9±2.1 (Table 6.2) to 3.5±1.8 mm.

One limitation of sCBCT is that it does not reproduce the streaking artefacts due to motion in CBCT imaging. When streaks were artificially added, the effect on the dCT accuracy was small: a difference of up to 0.3 mm in WET_{RMS} measured in the distal and proximal surfaces, and on the PTV; the DD_{RMS} value increased by 0.3 %mD.

Table 6.3: *Overall uncertainty in dose estimation. Comparison between rCT dose (recalculated on TPS) and pCT dose and range-corrected dose (based on the dCT): similarity between the 50% and 90% isodose lines and 3D dose distributions (mean value ± standard deviation). The results were averaged over all patients, fields, registration algorithms and CBCT datasets.*

	Method	DT_{mean} (mm)	DT_{RMS} (mm)	$DT_{95\%}$ (mm)	DSC
50% isodose	dCT	0.3±0.2	0.8±0.5	1.4±0.9	0.98±0.01
	pCT	0.5±0.3	1.1±0.8	2.2±2.0	0.97±0.01
90% isodose	dCT	0.3±0.1	0.7±0.2	1.4±0.4	0.96±0.05
	pCT	0.4±0.6	0.9±0.6	1.9±1.7	0.95±0.05
		DD_{mean} (%mD)	DD_{RMS} (%mD)	$DD_{95\%}$ (%mD)	Gamma-pass (3%/3mm)
3D dose	dCT	-0.2±0.9	6±2	12±6	86±5
	pCT	-0.1±2.6	10±5	21±16	85±5

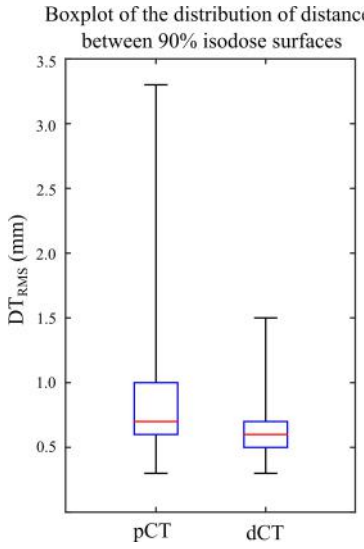

Figure 6.12: *Boxplot of the DT_{RMS} values found for the 90% isodose curve over the whole patient cohort using pCT and dCT (in comparison with rCT).*

Table 6.4: *Uncertainty in WET (dCT vs rCT) within the distal surface for the dCTs generated by each CBCT dataset, and corresponding comparison between rCT dose and range-corrected dose within the 3D dose distributions (mean value ± standard deviation). The results were averaged over all patients, fields and registration algorithms.*

	WET (distal surface)			Dose (3D dose)		
CBCT dataset	WET_{mean} (mm)	WET_{RMS} (mm)	$WET_{95\%}$ (mm)	DD_{mean} (%mD)	DD_{RMS} (%mD)	$DD_{95\%}$ (%mD)
rCBCT	0.0±3.0	4.8±2.3	10±4	-0.2±1.0	7±3	15±8
dCBCT	0.0±1.9	3.3±2.0	7±4	-0.4±1.0	6±2	11±6
sCBCT	1.6±1.7	3.7±1.7	7±4	-0.2±1.0	6±2	10±6

Table 6.5: *Uncertainty in WET (dCT vs rCT) within the distal surface for the dCTs generated by each DIR algorithm, and corresponding comparison between rCT dose and range-corrected dose within the 3D dose distributions (mean value ± standard deviation). The results were averaged over all patients, fields and CBCT datasets.*

	WET (distal surface)			Dose (3D dose)		
Registration algorithm	WET_{mean} (mm)	WET_{RMS} (mm)	$WET_{95\%}$ (mm)	DD_{mean} (%mD)	DD_{RMS} (%mD)	$DD_{95\%}$ (%mD)
REGGUI	0.9±2.0	3.6±1.8	7±4	-0.5±0.9	6±2	11±6
NiftyReg	0.7±2.5	4.1±2.1	8±4	-0.1±0.9	6±3	12±7
RayStation	-0.1±2.9	4.1±2.4	8±5	-0.2±1.0	7±3	13±7

6.3.1.3 Effect of different registration algorithms

The performance of the different DIR algorithms used to generate the dCTs is presented in Table 6.5. REGGUI performed slightly above the average; however the differences were generally small.

6.3.1.4 Effect of deformed CT correction

The dCT correction step was necessary for six out of the twenty cases (30%). These patients had the following anatomical changes that could not be recovered by DIR alone: atelectasis, lung reinflation, and complex tumour shrinkage (such as erosion from within the gross tumour and regression of infiltrating tumours). Figure 6.13 shows an example of an infiltrating tumour regression (top) and lung reinflation (bottom), where the correction step was necessary.

The impact of the correction step on WET and range-corrected dose uncertainty are presented in Table 6.6. Globally, the dCT correction step resulted in a reduction of the WET_{RMS} (distal edge) from 4.7±2.5 to 3.9±2.1 mm (Table 6.2). It also caused a reduction of DD_{RMS} of the 3D dose distribution from 7±3 to 6±3 %mD (Table 6.3). The correction algorithm improved the accuracy of the dCT for the six patients with substantial anatomical change; nevertheless, the overall uncertainties were still greater for these patients compared to the others.

6.3.1.5 Uncertainty due to the use of cone-beam CT for registration

Comparison between dCT_{rCT} and rCT provides an indication of the level of accuracy achievable with current registration algorithms. It further provides a quantitative assessment of the impact of DIR-based WET and dose calculations using CBCT. An additional uncertainty in WET of approximately 1 mm can be attributed to the poorer quality of the CBCT compared to CT (Table 6.7). However, image quality is not the only factor to explain

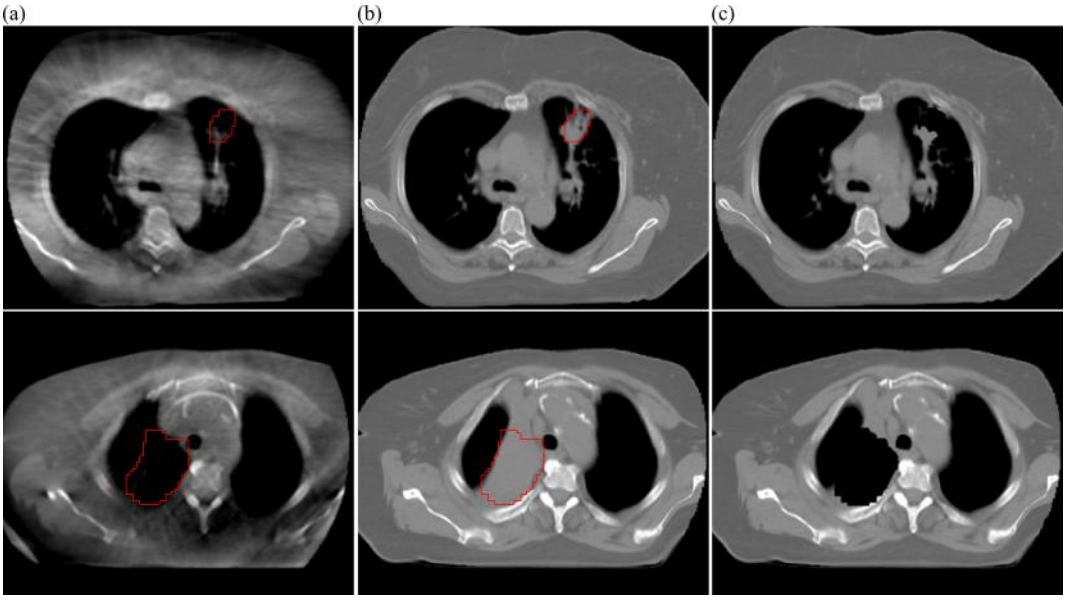

Figure 6.13: *Examples of the dCT correction: (a) rCBCT, (b) uncorrected dCT and (c) corrected dCT for a patient with tumour (top row, PT#8) and lung changes (bottom row, PT#1). The region defined by the red contour was identified by the correction algorithm as a region of mismatch between the dCT and CBCT and the HUs were replaced by bulk values of tissue/lung during the correction step. The dCTs were generated using REGGUI as registration algorithm.*

Table 6.6: *Uncertainty in WET (dCT vs rCT) within the distal surface for cases with and without the dCT correction step, and corresponding comparison between rCT dose and range-corrected dose within the 3D dose distributions (mean value ± standard deviation). The results were averaged over all patients, fields and CBCT datasets.*

		WET (distal surface)			Dose (3D dose)		
Patient Cohort	Correction applied	WET_{mean} (mm)	WET_{RMS} (mm)	$WET_{95\%}$ (mm)	DD_{mean} (%mD)	DD_{RMS} (%mD)	$DD_{95\%}$ (%mD)
A[a]	N	1±2	3.1±1.4	6±3	-0.4±0.9	6±2	11±6
B[b]	N	1±5	8±4	19±9	0.2±2.3	10±4	19±14
	Y	0±3	6±2	13±4	0.1±1.1	8±2	14±7
Cohort A[a]: No dCT correction necessary (70%)							
Cohort B[b]: dCT correction needed (30%)							

Table 6.7: *Uncertainty in WET (dCT vs rCT) within the distal surface for CT-to-CBCT and pCT-to-rCT registrations, and corresponding comparison between rescan and range-corrected dose within the 3D dose distributions (mean value ± standard deviation). The results were averaged over all patients, fields, and DIR algorithms (for pCT-to-rCT) and CBCT datasets (for pCT-to-CBCT).*

	WET (distal surface)			Dose (3D dose)		
Registration	WET_{mean} (mm)	WET_{RMS} (mm)	$WET_{95\%}$ (mm)	DD_{mean} (%mD)	DD_{RMS} (%mD)	$DD_{95\%}$ (%mD)
pCT-to-CBCT	0.5±2.4	3.9±2.1	8±4	-0.2±0.9	6±2	12±6
pCT-to-rCT	0.6±1.7	2.8±1.8	6±4	-0.3±0.7	5±2	8±4

the differences found between pCT-to-rCT and pCT-to-CBCT registrations; the variability in setup for the different CBCT datasets also contributes to the overall uncertainty.

6.3.1.6 Comparison of the deformed CT method to simpler methods

Table 6.8 shows the uncertainties in WET for the different methods. The dCT method resulted in the smallest uncertainties. The statistic WET_{mean} illustrates well the limitations of the rigidly-aligned pCT and intensity-corrected CBCT methods. The rigidly-aligned pCT method considers that the patient anatomy does not change, and therefore does not take in consideration tumour shrinkage; therefore, the method is biased toward overestimating the WET. The error measured is also indicative of the importance of monitoring the patients with regular imaging throughout the course of lung proton therapy. The intensity-corrected CBCT does not correct for scatter artefacts, that visually results in regions of lower intensity in the images; thus, the method is biased toward underestimating the WET. The bulk-density intensity assignment is less biased, and the associated uncertainties are closer to those of the dCT method. In spite of being in theory a simpler method, automatic segmentation methods have to be complex due to the low CBCT imaging quality; a viable but laborious alternative is manual delineation. Therefore bulk-density intensity assignment is a good alternative to dCT for cases where the correction setup is not enough to corrected for registration errors. The good results obtained for bulk-density intensity assignment versus intensity-corrected CBCT were the motivation behind using bulk-assignment in the dCT correction.

6.3.2 Clinical indicators of replanning

In appendix A a complete description of the results obtained for the variation in clinical indicators (Table A.1 for WET, and Table A.2 for dose), as well as a breakdown of all the correct predictions and false negatives/positives extracted from the clinical indicators per patient is provided (Table A.3). Within the next sections, the clinical indicators extracted for select case studies (i.e., individual patients) will be discussed, grouped per type of

Table 6.8: *Overall uncertainty in WET (vs rCT) within the planning target volume (PTV), and on the distal and proximal surfaces (mean value ± standard deviation), for the deformed CT (dCT), rigidly-aligned planning CT (pCT), bulk-density intensity assignment (BD) and intensity-corrected CBCT ($CBCT_c$). The results were averaged over all patients, fields, registration algorithms, and considering only the dCBCT dataset.*

Region of interest	Method	WET_{mean} (mm)	WET_{RMS} (mm)	$WET_{95\%}$ (mm)
Distal surface	dCT	0.0±1.9	3.3±2.0	7±4
	pCT	1±5	8±5	16±10
	BD	-1.8±2.2	4.4±1.5	8±3
	$CBCT_c$	-13±11	16±9	24±10
Proximal surface	dCT	-0.1±1.1	1.7±1.0	3.4±2.0
	pCT	0.0±2.1	3.2±1.5	6±3
	BD	0.0±1.4	2.2±1.2	4.3±2.1
	$CBCT_c$	-8±8	10±7	16±9
PTV	dCT	0.0±1.7	2.9±1.7	6±4
	pCT	1±4	6±3	13±8
	BD	-1.0±2.1	3.4±1.4	7±3
	$CBCT_c$	-11±9	13±8	20±9

anatomical changes; general discussion will follow this analysis.

6.3.2.1 Lung changes

Atelectasis is the collapse of lung that is sometimes reversible. PT#1 developed partial atelectasis at the upper left lobe during week two (Figure 6.14) resulting in increased WET along the beam paths and subsequent under-ranging for the two fields ($WET_{under\text{-}95\%}$=10.4/12.3 mm for LPO_1/LPO_2 field, Figure 6.15). Tumour coverage was compromised and a higher dose was delivered to the oesophagus (D_{max} from 50 Gy to 71/71/68 Gy for $D_{dCT\text{-}WET}$/D_{dCT}/D_{rCT}) which triggered immediate replanning. The dCT predicted similar dosimetric indicators as the rCT.

When tumours regress, the previously blocked airway can reopen and reinflate the collapsed lung (PT#3, Figure 6.16). Lung reinflation reduced the WET along the beam path, and caused beam over-ranging ($WET_{over\text{-}95\%}$=41.4/4.1 mm for RPO/PA field). The change in dose distribution compromised tumour coverage (iCTV $\Delta V_{99\%}$=-27/-27/-13% for $D_{dCT\text{-}WET}$/D_{dCT}/D_{rCT}), which triggered replanning. The predicted loss of coverage was higher in the dCT than in the rCT, which can be attributed to the partial truncation of the CBCT at the beam entrance (see section 6.3.2.2).

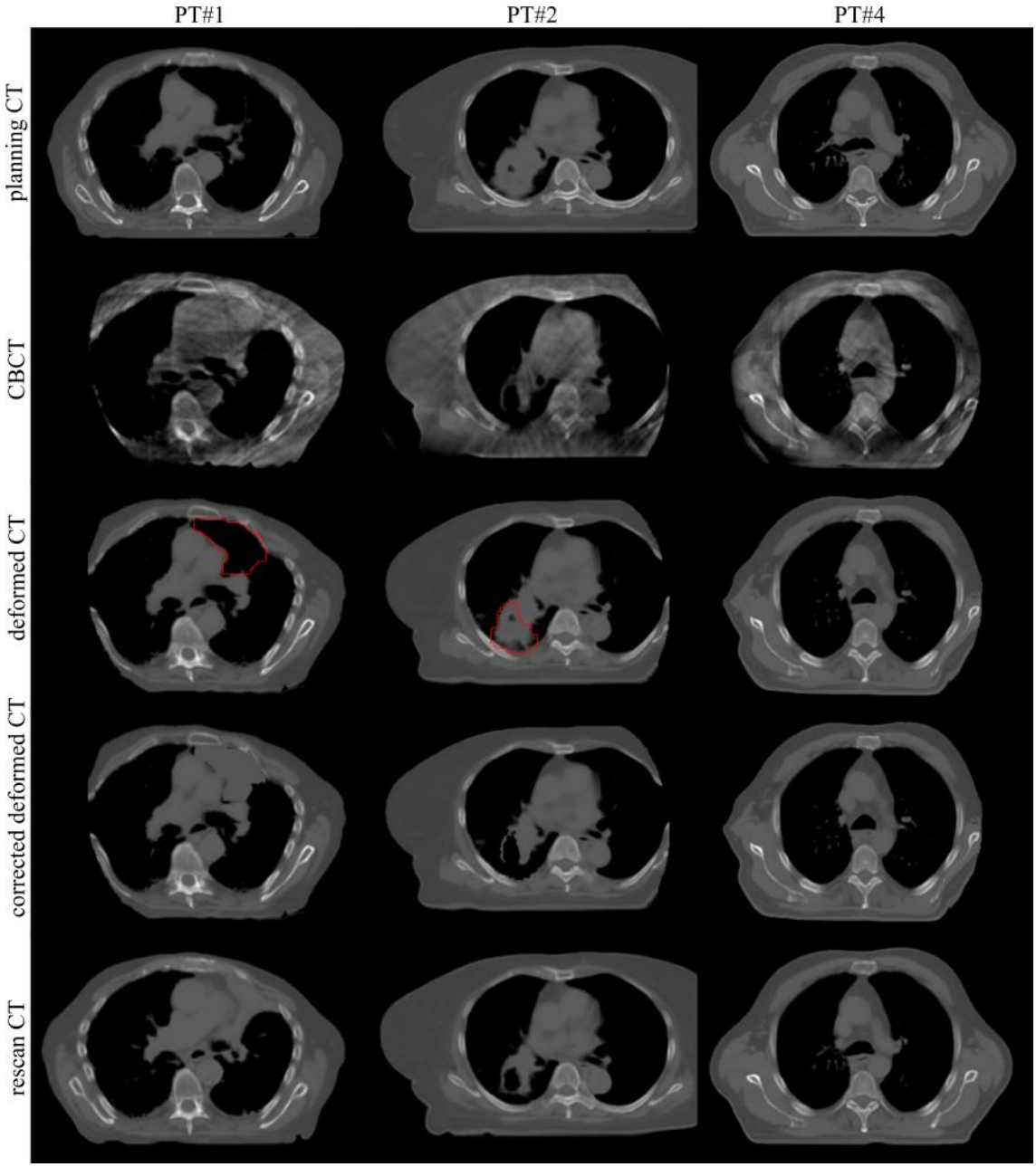

Figure 6.14: *Images used and generated by the workflow for PT#1 (atelectasis), PT#2 (large tumour shrinkage) and PT#4 (small tumour shrinkage). For PT#1 and PT#2, the dCT needed the correction step (region defined by the red contour). For PT#4, DIR alone recovered the changes well; however, the truncated CBCT data affected the similarity between dCT and rCT.*

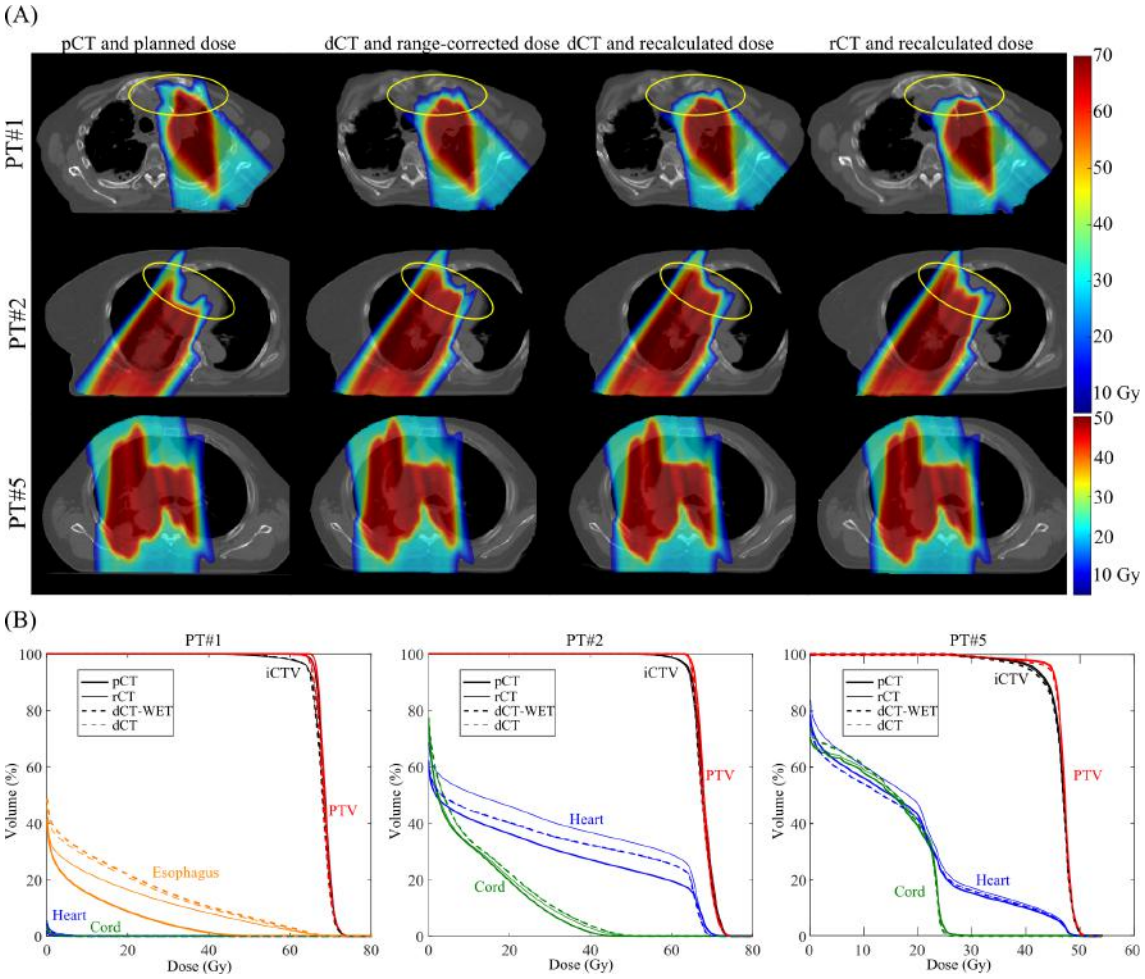

Figure 6.15: *(A) Color overlay of the CTs and corresponding dose distributions and (B) DVHs for PT#1, PT#2 and PT#5. For PT#1 the appearance of atelectasis increased the WET, resulting in under-ranging and loss of iCTV coverage. For PT#2 the shrinkage of the GTV resulted in decreased WET, and thus in over-ranging and increase in dose delivered to the heart and cord. For PT#5 there were only small changes in WET, and therefore all the dose distributions and DVH curves were similar.*

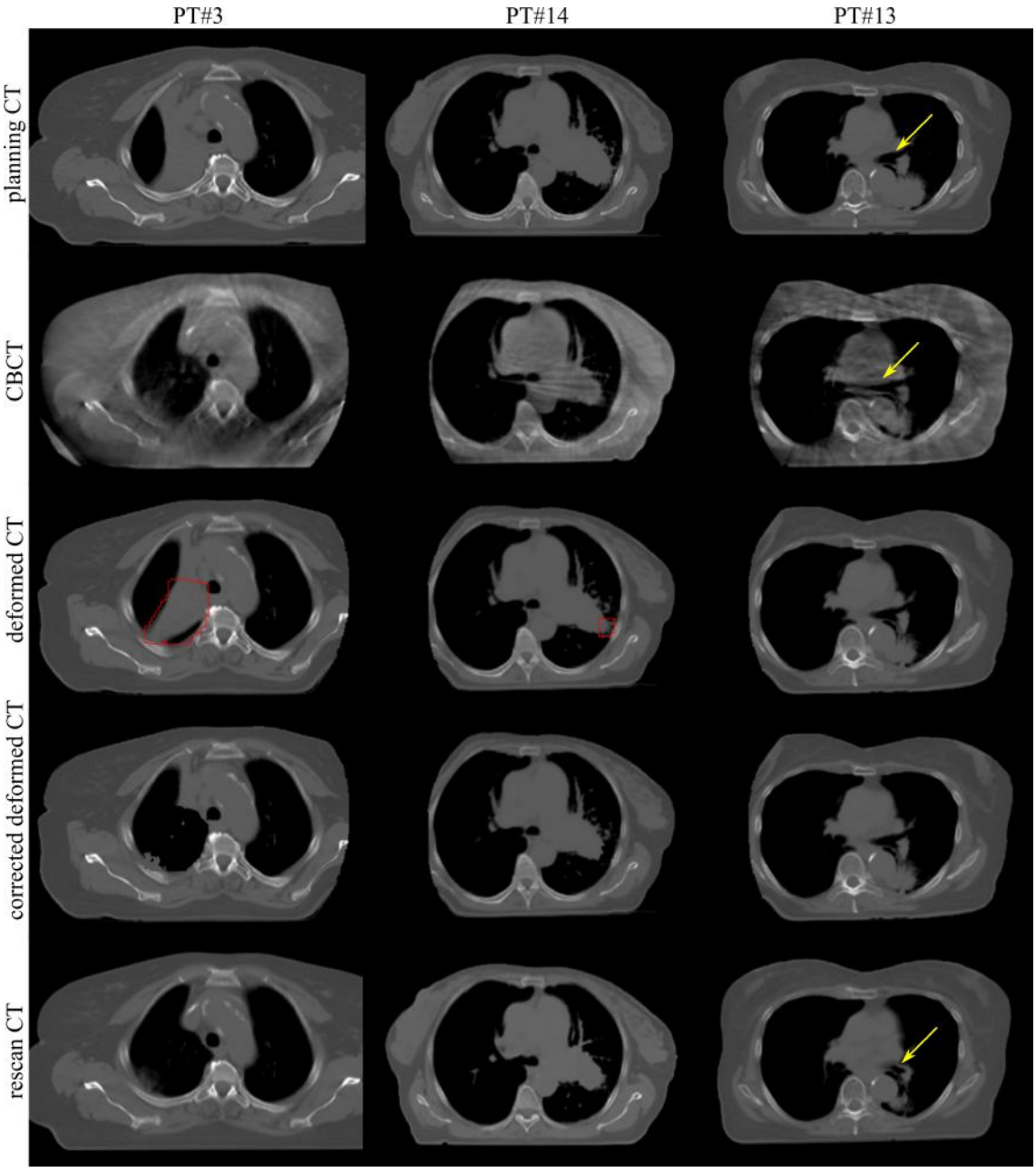

Figure 6.16: *Images used and generated by the workflow for PT#3 (lung reinflation), PT#14 (regression of infiltrating tumour) and PT#13 (shrinkage and changes in breathing pattern). For PT#3 and PT#14, correction of the dCT was necessary (region defined by the red contour). For PT#13, DIR was able to recover the tumour shrinkage; however, visible differences in setup occur between dCT and rCT, particularly in the relative position of the main airways (yellow arrow).*

6.3.2.2 Tumour changes

Different tumour response scenarios were identified and detailed below.

Infiltrating tumours For PT#14, the GTV decreased from 4.1 cm to 2.7 cm in diameter after four weeks of treatment (Figure 6.16). The uncorrected dCT resulted in DIR errors of the lung tissue between the tumour and chest wall. After applying the correction algorithm, the clinical indicators were nearly identical between dCT and rCT (iCTV $\Delta V_{99\%}$=-5/-7/-6% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$).

Tumour regression When tumours regress, the topological changes may not be handled by DIR alone. In PT#2 a 22.3 mm cavity appeared within the original tumour volume (Figure 6.14). Its size and location were accurately identified and accounted for by the dCT correction step. The reduction in WET along the beam path resulted in beam over-ranging to the heart ($WET_{over\text{-}95\%}$=24.6/25.2 mm for RPO_1/RPO_2 field, Figures 6.15 and 6.17). Dosimetric indicators between dCT and rCT were similar, i.e., decreased iCTV coverage (iCTV $\Delta V_{99\%}$=-7/-6/-8% at $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$), and increased dose to the cord (D_{max} from 45 to 52/52/49 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$) and to the heart ($V_{45Gy}$ from 25% to 31/31/35% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$, Figure 6.15).

Changes in tumour density PT#20 had both regression and changes in tumour density; the average intensity of the GTV decreased from 30 to -110 HU between the pCT/dCT and rCT, corresponding to a local WET variation of approximately 7 mm (Figure 6.18). The dCT retained the HUs from the pCT and underestimated the change in proton range, i.e., $WET_{over>3mm}$=27.0/40.7% for dCT/rCT (RPO_1 field). Regardless of the differences in WET between the dCT and rCT, identical reduction in dose coverage was detected in the DVHs.

Moderate shrinkage/enlargement Moderate tumour regression was here defined as a visually apparent change in tumour volume less than 25% of its original GTV. An example is PT#11, who has focal shrinkage (Figure 6.18) resulting in modest beam over-ranging ($WET_{over\text{-}95\%}$=8.5/7.5 mm for AP/RPO field), and an increase in dose to the cord (D_{max} from 30 to 36/37/37 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$). PT#16 was the only case of tumour enlargement during radiation treatment; the diameter increased 5 mm along the beam path (Figure 6.18). Because of the complex organ geometry at the mediastinum, both beam under and over-ranging were observed (i.e., $WET_{under\text{-}95\%}$=1.1 mm and $WET_{over\text{-}95\%}$=3.6 mm for the RPO field), resulting in increased dose to the cord (D_{max} from 35 to 46/47/40 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$), and a right shift of iCTV/PTV DVH curves. In these two cases the dCT and rCT offered similar clinical indicators with DIR alone (without the dCT correction step).

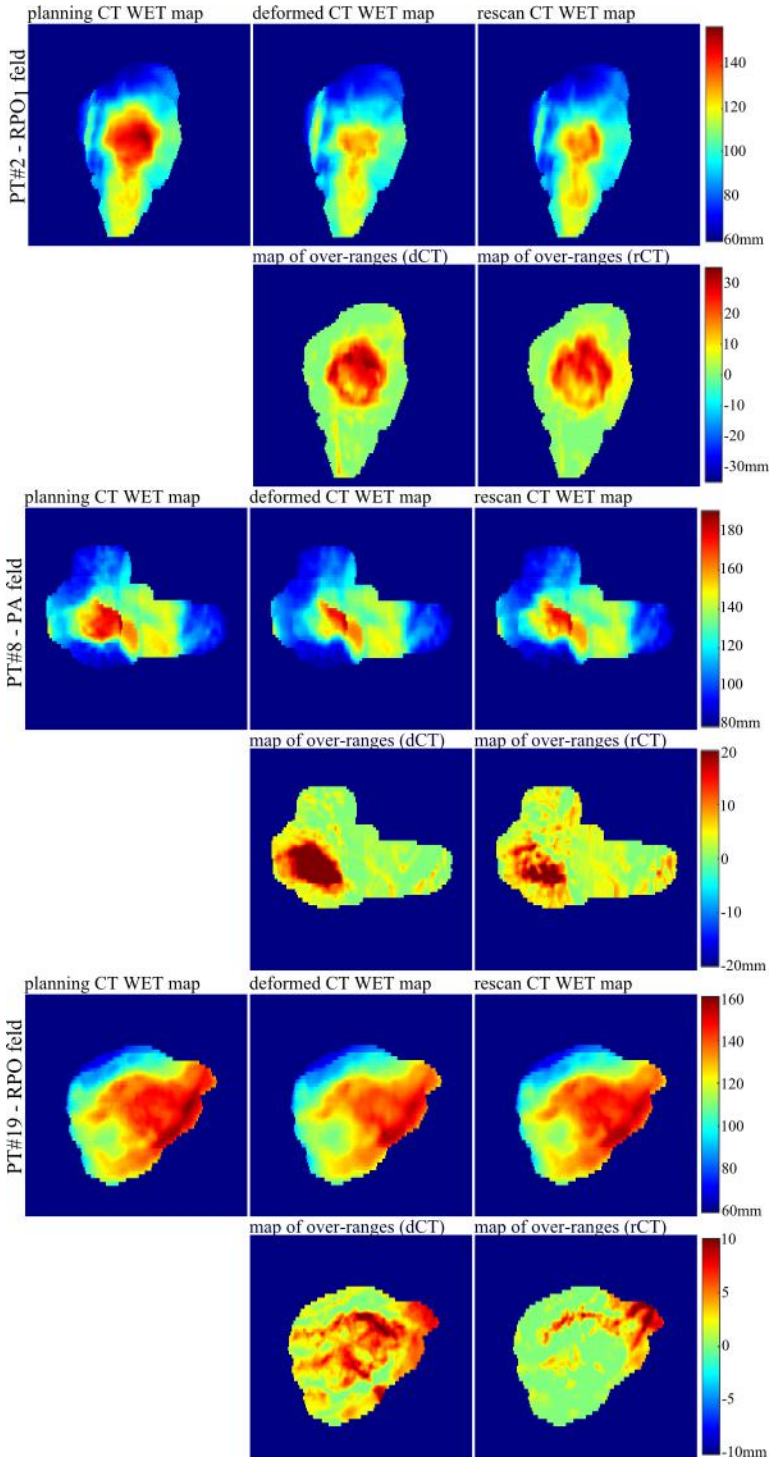

Figure 6.17: *WET maps for planning and deformed/rescan CT; and corresponding difference maps for PT#2 (RPO_1 field), PT#8 (PA field) and PT#19 (RPO field).*

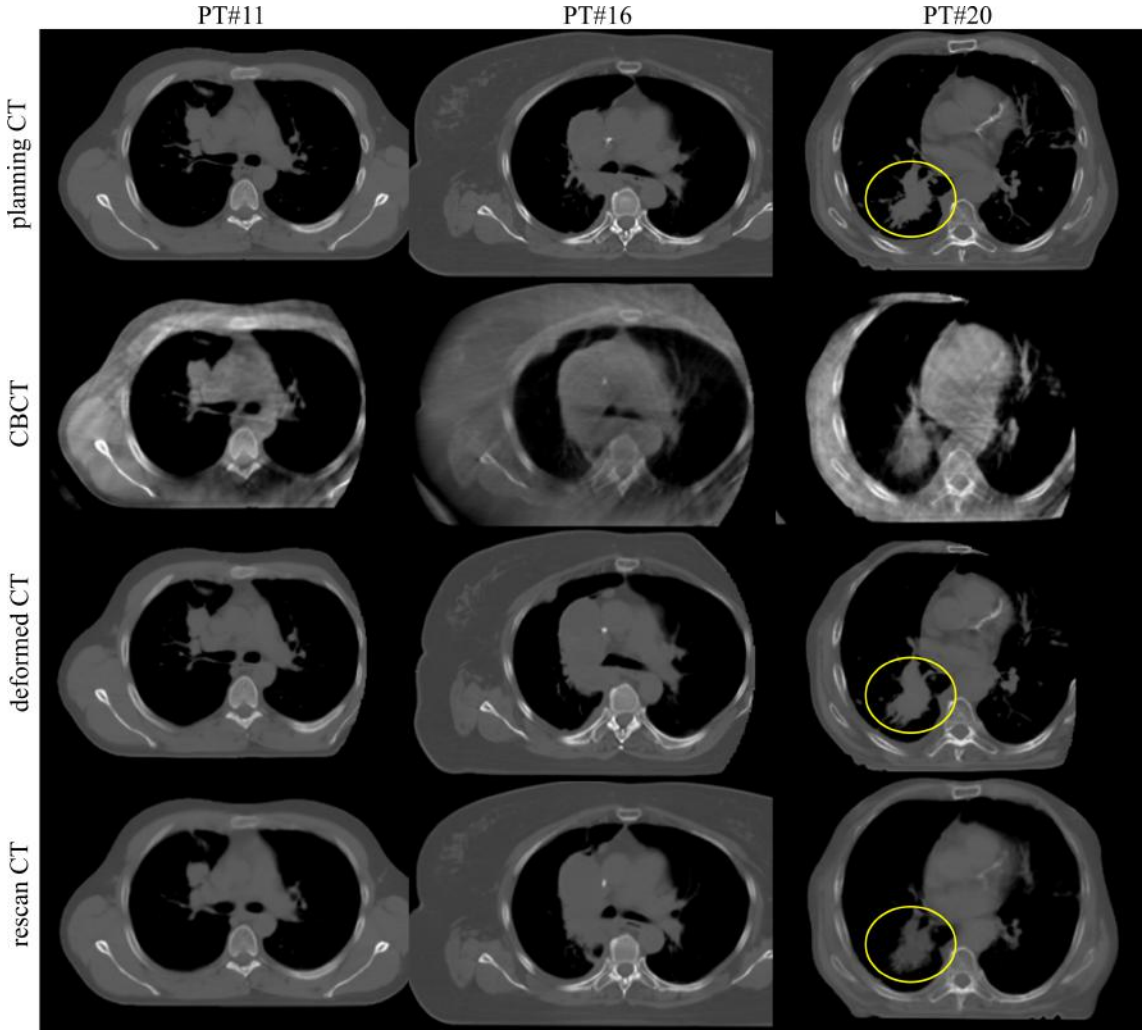

Figure 6.18: *Images used and generated by the workflow for PT#11 (focal shrinkage), PT#16 (moderate tumour enlargement) and PT#20 (tumour density changes). Only PT#20 required the dCT correction step, but not in the slice shown in this figure; the overall intensity in the gross tumour area (identified in yellow) changed from 30 HU to -110 HU.*

Other factors affecting WET and dosimetric indicators In addition to DIR errors, setup variations and differences in the respiratory pattern between CBCT and rCT scans can result in differences between $D_{dCT\text{-}WET}$/D_{dCT} and D_{rCT}. This implies that different clinical indicators are being extracted, and can, therefore, create false positive/negative triggers for offline review. The online workflow was less robust for patients in whom the magnitude of dose differences arising from setup variations was comparable to those arising from internal anatomical changes. For PT#13, tumour shrinkage was well recovered by DIR; however the position of the main bronchi was shifted superiorly in rCT in comparison to the rCBCT (Figure 6.16). The movement of the main airway cause different predictions of target coverage in different images (iCTV $\Delta V_{99\%}$=+3/+1/-4% for $D_{dCT\text{-}WET}$/D_{dCT}/D_{rCT}). WET difference maps were also affected: the magnitude of $WET_{over\text{-}95\%}$ was small but the 2D WET maps gave different indicators ($WET_{over>3mm}$ was 18.8/46.6% for dCT/rCT for PT#19 RPO field, see Figure 6.17). Clinically, none of these patients required replanning as the dosimetric changes were generally small and the effects averaged out during the course of treatment.

Setup errors should not be confused with systematic drift of tumour position through the treatment course. For PT#5, the primary tumour shifted in the inferior direction, and this was consistent between the CBCT and rCT. DIR accurately described the change in tumour position, but the modest change in WET had minimal effect on target coverage or dose to OAR (Figure 6.15).

Due to FoV limitations, a minority of CBCTs did not encompass the entire exterior of the patient body at the beam entrance (Figures 6.14, 6.16 and 6.18). If CBCT truncation is uncorrected, it may lead to inaccurate clinical indicators (PT#3: iCTV $\Delta V_{99\%}$=-27/-27/-13% for $D_{dCT\text{-}WET}$/D_{dCT}/D_{rCT} and, PT#17: iCTV $\Delta V_{99\%}$ =0/0/-5%).

6.3.2.3 General considerations

Eighteen of the twenty patients exhibited over-ranging of the proton beams. Two patients (PT#1 and PT#16) showed potential under-ranging. These were cases of atelectasis and tumour enlargement. For all patients, the average absolute difference in $WET_{over\text{-}95\%}$ and $WET_{over>3mm}$ between dCT and rCT were 3.4±2.7 mm and 12±12%, respectively. Figure 6.17 shows examples of WET and WET difference maps. WET difference maps identified the same regions of under/over-ranging for all patients with large anatomical changes. This was true even for PT#8 and PT#20, in spite of the full magnitude of the over-ranging not being fully recovered due to limitations of the corrected dCT to reproduce complex shrinkage and/or density changes. In cases of smaller changes or setup variations, the WET difference maps were less uniform. In general, values of $WET_{over\text{-}95\%}$ needed to exceed 15 mm before significant dosimetric changes could be detected. The most common issues that could lead to a replan were loss of tumour volume coverage, increase in maximum dose to the cord, and over-ranging of dose into the heart.

A 3D gamma-index method (3%/3mm) was used to compare the dose distributions using a 20% of the pD cut-off. Using D_{rCT} as gold-standard, the percentage of passing points for all patients were 88.5±6.1%, 89.4±4.7% and 90.1±4.4% for D_{pCT}, $D_{dCT\text{-}WET}$ and D_{dCT} respectively; when considering only patients with changes in lung (PT#1 and PT#3) and large changes in tumour volume (PT#2, PT#8, PT#14, PT#20) the values were 81.8±5.2%, 86.0±5.9% and 87.3±5.0%.

Considering the dosimetric statistics related with target coverage and OAR tolerances, the most common issues were loss of coverage and increase in dose to the heart and cord. For the remaining OAR investigated, in general, either the dose was quite low at planning stage and small variations were negligible, or the tolerances were never met due to PTV overlapping.

6.3.3 Discussion

CBCT plays an important role in IGRT. In proton therapy, quantitative applications require accurate HUs in order to make clinical decisions for ART. dCTs generated from CBCTs is one step in that direction and may play a complementary role to rCT. The careful application of CBCT to correct its deficiencies may permit it to replace the majority of verification scans for lung proton therapy. Quantifying its accuracy is therefore an important subject prior to clinical implementation.

The accuracy of dCT for WET estimation and consequently the "dose of the day", for lung cancer patients was investigated using multiple CBCT datasets and registration algorithms. Each type of CBCT has its own advantages as well as limitations, but taken together it offers enough evidence for quantifying the uncertainties associated with DIR and CBCT imaging for the proposed application. An overall uncertainty of 3.9±2.1 mm (RMS) in the WET was found at the PTV distal surface. The rCBCT resulted in the highest uncertainties due to discrepancies in positioning between rCBCT and rCT. The lower results obtained for the sCBCT were investigated. One specific region with larger DIR errors was in the posterior interface between lung, tissue and rib, where scatter artefacts degraded the contrast between the heterogenous tissues. In the sCBCT, scatter was overestimated in these regions, and was most likely the main source of discrepancy in WET calculation between the dCBCT and sCBCT datasets. The comparison between multiple registration algorithms provided a measure of the reproducibility of the method. Differences between algorithms were small and not significant when dose distributions were evaluated. This shows the versatility of the workflow proposed, which can easily be modified to be used with in-house or commercial DIR solutions available at other institutions. The ART workflow was implemented as a prototype in a research platform, and therefore not optimised for fast online use. However, the most computationally demanding processes (DIR, dCT correction and range-corrected dose) can be performed in several minutes using GPU.

DIR has inherent uncertainties and associated errors, especially at heterogeneous tissue interfaces, that lead to inaccuracies in WET and dose computation. The cohort of lung patients included in this study included a broad range of clinical situations that can compromise the accuracy of the DIR. The main limitation is related with the mathematical basis of the most common and popular DIR algorithms, whose deformations include only translation, expansion and contraction. Changes in topology, such as the appearance or disappearance of tissue, cannot be reproduced without introducing singularities in the deformation fields. Since the intensities are mapped from the CT to the CBCT space, variations in the density of the tissue between images cannot be reproduced [191]. In fact, subtle changes in density of lung tissues between planning and verification (PT#20) were undetectable on rCBCT but apparent on rCT. Additionally, the volume of rigid bodies is not necessarily preserved and the displacement between ribs and lung may not be accurately modelled without specialised algorithms [110, 192]. An important novel aspect of this work was the implementation of a correction step of the dCT. This correction step was a solution found to account for large changes in the lung, such as atelectasis (PT#1), lung reinflation (PT#3) and pleural effusion; and within the tumour, such as cases of tumour erosion (PT#2) or the regression of infiltrating tumours (PT#8, #14 and #20). This method improved the accuracy of the WET calculation and range-corrected dose estimation. For complex anatomical change such as pleural effusion, the appearance or disappearance of a small layer of fluid may be too thin to be identified by the dCT correction step. Further work is necessary to develop DIR algorithms (or additional corrections) that account for the remaining issues described above. Scatter artefacts in CBCT decreases image contrast and results in registration errors at the interface of lung-tissue-bone at the posterior rib wall. The low dose thorax setting with a half-scan was used for the CBCT acquisitions for all patients, and not adjusted for patient size; thus, larger patients had poorer images available. Improvement to image reconstruction is currently being worked by the vendor (IBA). Another common registration error was the positioning of the scapula, which can move in and out of the path of lateral oblique fields, and therefore cause variations in WET calculation. The CBCTs used in the study had a limited FoV which may result in a cropped external contour and introduce errors in the range-corrected dose for lateral oblique fields where the range was locally overestimated (4 out of 40 fields) due to missing tissue. On the current system design, the CBCT was retrofitted using the imaging panel for kV imaging. The panel was not designed to be offset laterally and this causes the current FoV limitation. This is not a limitation of the proposed workflow when CBCT systems with larger FoVs become available in the proton clinic; additionally this limitation can be mitigated by positioning the patient to capture the beam entrance during CBCT acquisition.

From a clinical perspective, two scenarios are possible with DIR errors. The first is a false positive trigger, i.e., the dose calculated on the dCT indicated a change in dosimetry when there is none. The outcome is an unnecessary CT scan to confirm the findings. The other is a false negative trigger, i.e., the dose calculated on the dCT failed to detect the

change in dosimetry. While this scenario poses a bigger risk it is unlikely to occur. Higher DIR errors are associated with larger anatomical changes, and in such cases variations in dosimetry are usually still predicted even if with a different magnitude.

Proton plans that are more robust to registration errors may be used to minimise the issues generated by the anatomical changes. Thus, along with the impact of CBCT datasets and DIR algorithms in the overall uncertainties reported, the impact of field direction was also investigated. Each patient was treated with two treatment fields, with the posterior-anterior (PA) direction being the most common treatment approach (85% of the fields). Within those, lateral oblique beam were used in 38% of the cases. Due to the distortions caused by the couch and cropped FoV, it was empirically expected to have larger errors in WET for lateral oblique fields. In fact, a larger uncertainty in WET was found for lateral oblique field than for posterior fields (WET_{RMS} of 3.4±1.6 vs 4.5±2.4 mm); however we could not definitely link the source of those differences to the beam direction. In fact, 42% of the lateral oblique fields corresponded to patients that required dCT correction against only 23% for posterior fields. Since a greater uncertainty occurs when dCT correction is required, this could be the reason of the differences found. A larger patient cohort is necessary to further investigate the impact of beam direction.

Using dCT directly for dose recalculation is a viable alternative to remove errors associated with DIR. Despite the vast work on directly using CBCT in conventional photon treatments [9, 97, 99], its usability is still limited in proton therapy, as was shown in chapter 5 and in other studies [193], and the dCT is a viable interim solution. Other groups are working on more elegant approaches to improve the HU accuracy [194, 195]. The direct use of CBCT leaves other problems unanswered, such as limitations of the FoV and contour propagation. Other approaches to estimate the "dose of the day" should also be investigated; while range-corrected dose can be computed rapidly it is still only an approximation. The uncertainty of the "dose of the day" computed online may be reduced for instance by incorporating in the workflow a fast GPU-Monte Carlo calculation [196].

One of the unexplored topics at the current stage of this work was the evaluation of the workflow for automatic segmentation. The similarity between manually delineated and automatically propagated (using the deformation fields) structures is a popular method to validate DIR for clinical applications [10]. The open-source algorithms, NiftyReg (chapter 3) and REGGUI [11, 12], had been previously validated in the context of HN malignancies, and therefore their appropriateness for a similar clinical application had been already investigated. Additionally, fast segmentation is a prerequisite the clinical translation of a fast online ART workflow. Uncertainty in segmentation propagates to dose-based clinical indicators of the need of replanning (for example, maximum dose to the spinal chord), and this is a topic to be investigated in the future. Additionally, the impact of the dCT correction step has to be accounted for when propagating affected structures.

Regarding the investigation of clinical indicators extracted from the dCT, although the dCT may not reproduce identical WET maps, it identifies the same trends as the rCT regarding the effect of the WET changes. 90% of the fields with $WET_{under\text{-}95\%}/WET_{over\text{-}95\%}$ larger than 10 mm were properly identified as such from the dCT. The dose mapping method reproduced similar clinical indicators for patients with considerable changes that may trigger a replan. The most common issue was loss of target coverage of the iCTV. For PT#1, PT#2 and PT#14, the impact to OARs was detected (oesophagus, heart/cord and cord respectively). For PT#8, the changes in OAR dose were not properly detected, while for PT#20 an increase in cord dose was incorrectly detected. When smaller changes occurred, differences in OAR dose were also detected (PT#11, PT#12 and PT#16), but some false positives/negatives occurred for loss of target coverage (PT#13, PT#15 and PT#17). Variations in setup can result in overestimation (PT#9) and underestimation (PT#19) of over-ranging, but with minimal dosimetric impact. In general, OAR doses were maintained within tolerance; however, special care should be given to fields that point towards an OAR, such as lateral oblique fields that may range out at the heart (PT#2, PT#7 and PT#12).

An important conclusion taken from this retrospective investigation was the necessity to evaluate multiple parameters during the decision-making process: changes in WET, quantitative review of images and dose distributions, DVHs and corresponding dose statistics. Flags raised by a single indicator should be backed up by additional evidence. For example, the iCTV $V_{99\%}$ statistic was quite sensitive even when the DVHs did not reflect major changes. Similarly, large discrepancy in $WET_{over>3mm}/WET_{under>3mm}$ could be found while the 2D WET maps were consistent in terms of identified areas of considerable over/under-ranging. In cases where the anatomical changes are small the decision to replan should not be based on individual scans, but rather on continued monitoring. Smaller changes can in fact be comparable to setup errors, and may average out.

The rCBCT and rCT were acquired close in time but were not identical due to setup errors. This is not ideal to validate the clinical workflow and indicators, but is representative of what would happen on prospective patients (i.e., a rCT would be ordered based on the CBCT and setup variation cannot be avoided). Daily positioning variations, such relative position of the trachea (PT#13), pacemaker wires (PT#7) and shape of external contours can influence the range of the proton beams. This also highlights one of the limitations of this study, which was to look at a single time point per patient. At the moment not enough data was available to evaluate multiple images per patient. The promising results of this first study will be motivation to continue acquiring CBCT and rCT data simultaneously.

The range-corrected dose distributions lead to very similar clinical indicators as the recalculated doses on the TPS, and in general similar dosimetric indicators were extracted. This intermediate step in a clinical ART workflow is of utmost importance to carefully review all the clinical indicators and avoid requesting unnecessary verification scans due

to gross operator errors. It is also the most opportune moment to manually edit the dCT in case gross registration errors were not properly accounting for by the correction step. The translation of the workflow to prospective clinical routine will allow to build confidence, and possibly future simplification of the workflow.

CBCT has just recently become available at one of the PSPT treatment rooms at HUP, hence the reason for this first clinical application to be performed in lung patients. The method can be directly applicable to spot scanning treatments, as the "dose of the day" is estimated based on previously calculated dose distributions and WET maps only, and not on plan-specific parameters. However, compensator smearing is not used in spot scanning; thus, the plan robustness has to be assessed differently. It can also be fine-tuned for other treatment sites. Currently, the plan is to extend the CBCT system to two additional gantries at HUP in 2016; therefore within the next months this study will be extended to HN malignancies. CBCT imaging has an important role in image-guided proton therapy, particularly as a substitute for verification scans. In addition to contribute to patient outcome by allowing to identify those that benefit from replanning and reducing the dose delivered in imaging, there are other important benefits from an operational perspective. Depending on the frequency of scans and volume of patients, verification CT scans take up significant resources and time on the CT scanners. Substitution of the majority of verification scans with CBCT will have positive operational impact for the clinic.

Currently the decision support for replanning is based on dosimetric analysis of the rCT, using dose clinical indicators used in this study. The set of clinical indicators investigated here were thus chosen as an empirical and sensible starting point for replanning threshold levels, but were not fully optimised by any means. Based on the data gathered, large values of under/over-ranges ($WET_{under\text{-}95\%}$ >10 mm and $WET_{over\text{-}95\%}$ >15 mm), dose to the heart/cord (ΔV_{45Gy} >15% and ΔD_{max} >5 Gy), shifts of the DVH curves to lower doses for targets (PTV/iCTV) are good indicators and thresholds to trigger offline review. To develop an optimised decision support for replanning was out of the scope of this project, but is a natural evolution of the work here presented. It is crucial for clinical translation to identify the most adequate clinical indicators, and define clear thresholds that trigger the following action-level up until the replanning decision. On the technical side, improvement of the workflow to minimise its current limitations is a priority, which includes investigating lung-specialised DIR algorithms, automatic segmentation validation, improvement of CBCT image quality and integration with TPS or using more accurate dose calculations.

6.3.4 Conclusions

A CBCT and DIR based adaptive proton therapy workflow for lung cancer patients was proposed and evaluated. In the workflow a fast calculation of the "dose of the

day" was implemented by computing the WET along the beam paths in the corrected dCTs. The accuracy of the method was benchmarked to the forward dose calculations using the rCTs. In addition, multiple surrogates of an ideal CBCT were created to assess the accuracy of the calculations; different DIR algorithms were applied to estimate the reproducibility of the results. The dCT provided similar WET and dosimetric information as a rCT in a multitude of clinical scenarios. However, one important finding was that DIR alone could not fully reproduce all of the complex changes in the thorax and a novel correction step was proposed to deal with gross registration failures.

Chapter 7

Multimodal and multitemporal imaging in radiotherapy

> If a man never contradicts himself, the reason must be that he virtually never says anything at all.
>
> Erwin Schrödinger

With the advent of functional imaging to clinical settings it becomes important to incorporate the information from various imaging modalities (particularly MRI) into the radiotherapy pathway. The main objective of this chapter was to investigate the feasibility of automatic DIR in registering multimodal and multitemporal CT and MR images. This study was, as in chapters 2 to 5, focused on HN malignancies.

The work in this chapter resulted in the following output:

- C. Veiga, R. Mendes, D. Kittapa, S.-L. Wong, R. Bodey, M. Modat, S. Ourselin, G. Royle, and J. McClelland, "Optimization of Multimodal and Multitemporal Deformable Image Registration for Head and Neck Cancer", Imaging and Computer Assistance in Radiation Therapy Workshop of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention (Munich, Germany, 2015).

7.1 The role of multimodal and multiparametric imaging in radiotherapy

The most widespread concept of treatment adaptation in radiotherapy is to modify, in an online and automated fashion, the treatment being delivered such that the physical dose delivered matches the dose originally planned, considering the possible temporal

anatomical changes of the patient throughout the treatment course. These changes may be tumour shrinkage and weight loss during the course of the treatment, such as the clinical problems studied in the previous chapters. However this kind of approach only considers physical dose maps and ignores radiobiological aspects; in fact, the relationship between biological effect and physical dose is not linear. When assessing and verifying the success of a treatment, biological effects and final outcome (e.g: cure or toxicity) are more relevant than the actual physical dose map. However the aim is, and has always been, to obtain a dose distribution that based on clinical experience and biological knowledge available, will result in the best clinical outcome. Therefore, treatment adaptation should also be guided by biological parameters [197].

Biological parameters can potentially be measured using sequential multimodal and functional imaging techniques. Functional imaging stands for imaging modalities that measure quantitatively biological parameters in healthy tissues and tumours. The pixel intensity of the images produced can be directly related with a particular physical, chemical or biological property of the tissue. MRI, magnetic resonance spectroscopy (MRS), single-photon emission computed tomography (SPECT), PET and ultrasound offer additional and complementary information of the tumour at the physiological, cellular and molecular levels. Such imaging techniques can provide non-invasive, quantitative and 3D characterisation of diverse physiological mechanisms, such as blood flow, vessel permeability, cellularity, cell membrane turnover, cell surface receptor expression, apoptosis, metabolism, cell proliferation and hypoxia [198]. Potentially, this means being able to quantitatively measure the tumour growth, tumour status and predict treatment response and outcome early on. This early assessment would allow for optimisation of individualised treatment for patients with the final objective of maximising treatment outcome, reducing toxicity and morbidity [199]. This additional information can be incorporated in an ART workflow to assess early the outcome of a treatment, and intervene if the patient response suggests that adaptation is necessary. Introducing additional multimodal and functional imaging into the patient pathway has promising applications in radiotherapy. The different areas of interest are summarised below:

Pre-treatment imaging

- Pre-treatment prediction of tumour response and personalised medicine: Radiomics, data mining of large patients databases and machine learning can be used to generate models to predict treatment outcome [200]. Additional imaging can generate more accurate models and is key for the widespread of personalised medicine. For example, when the cancer is particularly radio-resistant, more aggressive therapies may be considered.
- Treatment planning strategies: Additional imaging has been used in several applications to treatment planning such as (i) target delineation [59], (ii) dose boost to radio-resistant regions of the tumour (also known as dose painting [201]), and (iii) to

decide on the choice of beam arrangements to avoid regions of high radio-sensitivity or functionality.

Imaging throughout the course of treatment, including image-guidance

- Predicting response early after initiating therapy: The use of some functional imaging has been shown to predict the success of therapy before conventional measurements of size are measurable [202]. Anatomical and morphological changes often occur temporally downstream from the underlying physiological, cellular and molecular changes, so appropriate imaging metrics may be able to provide insight into early disease response and progression [198]. The early recognition of failure to respond to a specific treatment may allow alternative treatments to be explored, and therefore avoiding unnecessary radiation exposure and associated side effects [203].

- Treatment adaptation: Imaging information can be used to feed an ART workflow, by providing complementary information on anatomical and functional modifications of the targets (such as tumour shrinkage or modification in its subregions of elevated radio-resistance). This is of particular interest for MR-LINAC systems, where MR is used for IGRT, and therefore available at the moment of treatment [204, 205].

- Improved understanding of the biologic effects of therapy: Measuring the temporal biological changes can infer how physical dose correlates with biological effect, and track the mechanisms and events that lead to the success/failure of the therapy.

Pos-treatment imaging

- Assessment of treatment response: For many years the standard way to assess the patient's response to treatment has been to measure tumour size on longitudinal CT or MR scans, using 1D and/or 2D criteria, such as World Health Organization (WHO) guidelines [206] and the response evaluation criteria for solid tumour (RECIST) [207], respectively. Such anatomy-based response-assessment techniques have inherent limitations because 1D and/or 2D measurements are used to quantify changes in 3D. Arbitrary cut-off values categorise response and progression, do not take in consideration changes in tumour density, and cannot distinguish viable from dead tumour regions [208]. The increasingly recognition of the importance of "beyond anatomy" imaging is proved in the 2000 amendment of RECIST [209], which includes FDG-PET in assessing progressive disease [198]. Additional imaging allows for a more complete assessment of treatment response [210], and opens up the possibility of correlating physical dose and biological properties with the outcome.

7.2 Rationale

The idea of incorporating biological information in the radiotherapy pathway is not *per se* novel. As a matter of fact, several strategies are part of routine clinical work, such as dose boost, dose escalation and lower-dose irradiation of structures suspected of infiltration but without clear signs of disease. The big challenge is how to incorporate increasingly amounts of biological information relevant to radiotherapy response into planning, adaptation and follow-up of individual patients. The ability to make firm conclusion related with the usefulness of defining sub-regions of interest in the patient anatomy with anatomical and functional imaging method relies on the accuracy to co-register multiple sources of information [211]. Therefore, accurate image registration is a key part in the different applications of multimodal and multiparametric imaging. In the HN region, MR images are harder to register than other modalities (such as CT and CBCT) for several reasons: (i) the image resolution is poorer, particularly in the shoulders and thoracic area (ii) lack of proper immobilisation (i.e., treatment positioning may not be possible in the MRI scanner due to the coils and patient comfort) [212], (iii) image-specific artefacts, such as those caused by the inhomogeneities of the magnetic field (bias), and (iv) limited FoV of the current clinical acquisition protocols.

Due to the difficulties in registering these datasets most of the work done in MR registration uses rigid alignment or semi-automatic DIR guided by manually annotated landmarks. Not many studies have been done on validating automatic DIR for multimodal and multitemporal data in the HN region in the context of radiotherapy. Leibfarth et al. use DIR between planning PET/MR and CT images for HN patients, comparing three different optimisation metrics of a B-Spline DIR for dose painting applications [213]. Slagmolen et al. present a small feasibility study on CT-to-MR and MR-to-MR DIR for radiotherapy treatment planning [214]. On a more technical side, some groups have been developing specialised DIR algorithms for multimodal imaging [215, 216]. Other authors looked a various applications in different anatomical sites [217–219].

In this work the use of DIR for multimodal and multitemporal registrations was investigated and optimised. The stationary velocity fields implementation available in NiftyReg is used to co-register images from different modalities at similar timepoints (CT and MR) and from the same modality at different timepoints (MR) to the same reference space, and the quality of the registrations was assessed using manually annotated structures and properties of the DVFs.

7.3 Methods and materials

7.3.1 Patient data acquisition

The data used in this study was selected from a readily available database of HN patients acquired in a pre-existing and on-going clinical trial at UCLH. The aim of the trial was to ascertain the sensitivity and specificity of multiparametric MR for the detection of active disease in post chemoradiation residual tissue masses, and to investigate the prognostic value of conventional anatomical and functional MRI in determining early and final treatment outcome in HN patients. Therefore, each patient included in the trial would receive a routine radiotherapy pCT, a pre-treatment MR in treatment position booked as close as possible in time to the pCT (MR_1), and a follow-up MR six months after treatment (MR_2). The MR sequences acquired consisted of T_1- and T_2-weighted, T_1-fat saturated, diffusion-weighted (DW), blood oxygen level dependent (BOLD), and dynamic contrast enhancement (DCE). Different patients had different combinations of multiparametric imaging acquired, and the acquisition mode common between all datasets was T_2-weighted. A total of approximately 40 patients had completed the trial at the time of this study.

The clinical data available had a considerable amount of anatomical and functional information available, which in theory allowed for potential future studies built on the work here described. However, the aims of the clinical trial were not defined with the application of DIR here proposed in mind, and therefore the imaging acquisition protocols were clearly sub-optimal for our study. Clipping of the FoV of MRI images was a common acquisition problem that affects the registrations. A complete body contour is crucial for global alignment and to avoid registration errors near the edge of the FoV. Additionally, some patients did not acquire the first MR as they could not endure being inside the MRI scanner. For other patients the images were acquired but not in treatment position. In these cases (and similarly to pos-treatment images) the pCT and MR_1 had very different positioning, particularly in the flexion of the neck and lack of neck support system. This variability in the data required selection of adequate datasets and grouping based on differences in pre-treatment setup.

Thus, a total of 8 datasets were used to qualitatively investigate the most appropriate strategies for DIR; however, only 3 out the 8 datasets were used for in-depth quantitative analysis. The remaining patients were not included at this stage due to incomplete expert delineations. The inclusion criteria of this study was solely based on minimising acquisition issues characteristic of routine MR (such as clipping of the FoV), and not to select patients with smaller anatomical changes (Figures 7.1 and 7.2). These 8 datasets were all cases of MR_1 acquired in treatment position.

The pCT imaging protocol was as described in section 2.5.1.1. The MR images were acquired using the MAGNETOM Avanto (Siemens Healthcare, Erlangen, Germany) MRI

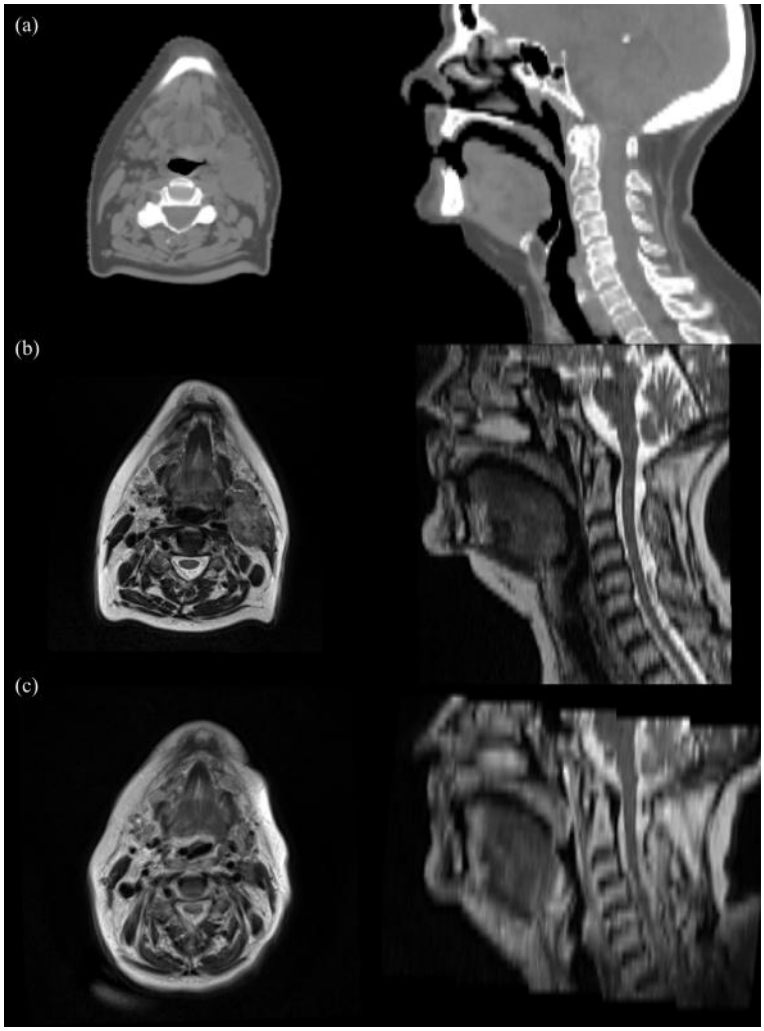

Figure 7.1: *Example of a patient included in this study: axial (left) and sagittal (right) view of (a) planning CT, (b) pre-treatment and (c) pos-treatment MRs. The pos-treatment MR was not acquired in treatment position.*

scanner (1.5T). In T_2-weighted images, TE varied between 90 and 110 ms, TR between 2400 and 8100 ms, slice thickness between 3 or 5 mm with a gap of 0.5 mm, and number of slices between 29 and 61. In addition, image resolution was 0.703×0.703 mm^2 or 0.859×0.859 mm^2.

7.3.2 Multimodal and multiparametric imaging in a radiotherapy workflow

For integration of MR data into the radiotherapy pathway, image registration is necessary between CT at the planning stage and repeat MR at different time points. To co-register multitemporal MRs with CT, two registration pathways can be followed (Figure 7.3):

1. the pos-treatment MR is registered with the pre-treatment MR, which is independently registered to the CT;

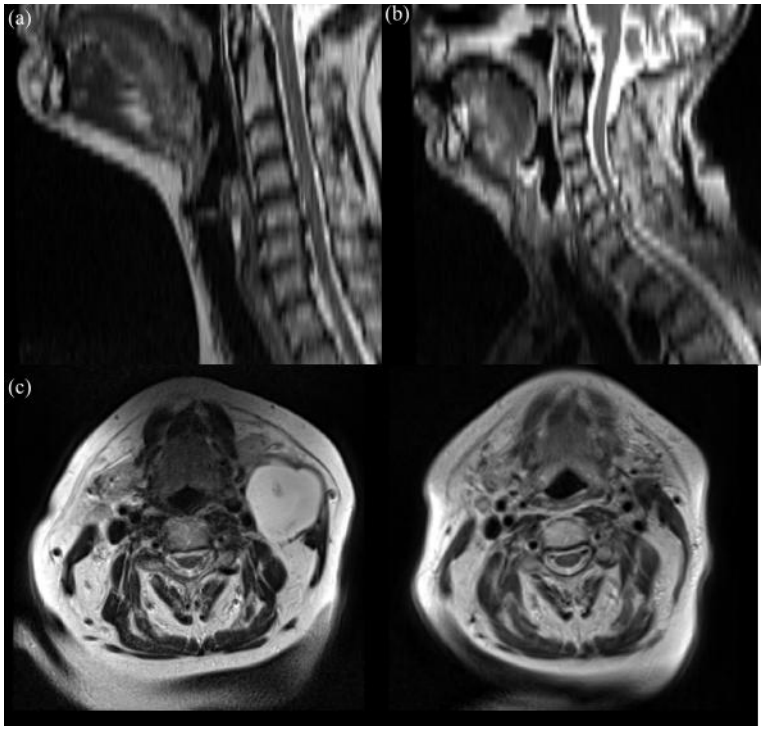

Figure 7.2: *Example of limitations of the MR data available: sagittal view of (a) reduced field-of-view issues and (b) poor image quality and (c) large anatomical differences between pre-treatment (left) and pos-treatment (right) MRs, such as patient weight and tumour volume (for this particular patient the pre-treatment MR was not acquired in treatment position).*

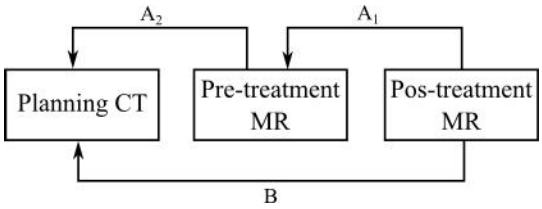

Figure 7.3: *Schematic diagram of registration pathways. To register CT with a pos-treatment MR two pathways can be followed: (A) the pos-treatment MR is registered with the pre-MR, which is registered to the CT or (B) the pos-treatment MR is registered directly with the CT.*

2. the pos-treatment MR is registered directly with the CT.

In this work only the results from the first pathway were assessed quantitatively. Due to the 6 months gap between pCT and MR_2 it was very challenging to tune the DIR parameters to be universally good, and in general the results were poor and physically implausible. Therefore, it was a preferable approach to independently register similar anatomical information from different modalities at similar time points (pCT-MR_1) and anatomical deformations from the same modalities at different time points (MR_1-MR_2). This allows to decouple the difference in image intensity between modalities from the anatomical deformations that occur over time. Therefore, two registration methods were investigated:

- CT-MR_1: if the two images were acquired close in time and with same immobilisa-

tion, a rigid registration is the easier and natural approach. However, DIR can be used to compensate for residual setup errors. This may introduce additional issues, which will be investigated here.

- MR_1-MR_2: monomodal DIR was investigated to track over time changes in anatomy. The ability to map anatomy between time points also allows to propagate co-registered functional information (using the same DVFs as in the anatomical sequences).

7.3.3 Image registration settings

NiftyReg stationary velocity fields implementation was the algorithm chosen for the registrations. As seen in chapter 4 it generates registrations with a geometric matching comparable to the standard unidirectional and asymmetric algorithms, while ensuring more desirable physical properties such as symmetry and inverse-consistency. DVFs are generated in both registration direction, facilitating both contour propagation between modalities and time points. Additionally, to propagate functional information acquired simultaneously with the anatomical sequences theoretically more plausible registrations are preferable since it involves tracking voxel-by-voxel changes over time of quantitative biological markers, such as the apparent diffusion coefficient (ADC) maps from DW-MRI.

Preliminary evaluation of the DIR quality was performed to find suitable parameters for the registrations, similarly to the work described in section 2.4 for CT-to-CBCT registrations. Several combinations of weight of the penalty terms, CPS and registration strategies were investigated. A total of six registrations parameters were selected and now more extensively tested per application. NMI was chosen as similarity measure for multimodal registrations, and LNCC for monomodal registrations. LNCC was preferred over other monomodal similarity measures since it handles better the non-uniform biases that cause artefacts in MR images [111]. Unlike the CT-to-CBCT registrations, here the CPS was defined in mm due to the varying resolution between datasets. Additionally, since the tumour is visible in some images (pCT and MR_1) and not in others (MR_2) the tumour was masked out to avoid optimisation of the registration in regions were there is no anatomical matching. CT-MR_1 registrations with immobilisation require larger constrains to reduce the risk of the registration causing additional uncertainties in comparison with rigid-only alignment (such as deformation of bones, etc). When the immobilisation is not present, those constrains are relaxed to give the algorithm enough freedom to recover larger deformations.

A total of six registration parameters with variable weight of the BE and CPS were investigated per registration type. The values of BE varied between 0.01% and 1%, while the CPS values tested ranged from 8 to 12 mm. NMI was chosen as similarity measure for multimodal registrations, and LNCC for monomodal registrations. LNCC was preferred over other monomodal similarity measures since it handles better the non-uniform biases

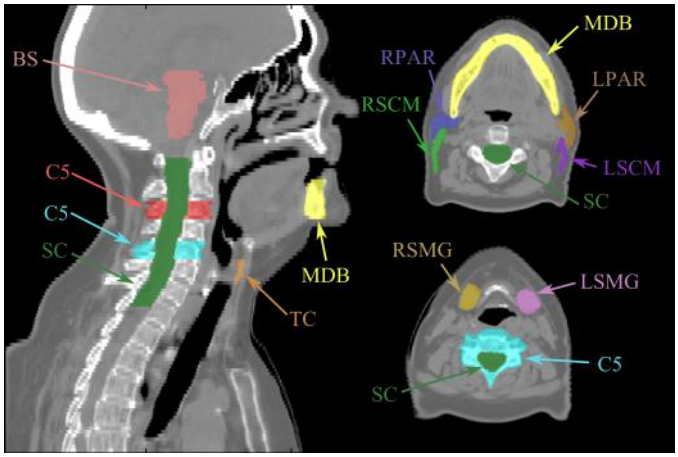

Figure 7.4: *Structure set manually delineated on the CT and MRs of each patient. This set consisted of vertebrae C3 and C5, mandible, thyroid cartilage (bony anatomy), spinal canal, brainstem, parotids (OARs), submandibular gland and sternocleidomastoid muscles (soft tissues).*

that cause artefacts in MR images [111]. To minimise the impact of artefacts due to field inhomogeneities, the MR images were corrected for bias using the N4ITK algorithm [220], which is available incorporated in NifTK. To avoid the optimisation of the transformation in regions where there is no anatomical matching, the tumour was masked out in MR_1-MR_2 registrations for patients where the gross tumour disappeared between MR_1 and MR_2 (as a result of the treatment). This avoids unrealistic deformations in these regions of no real one-to-one matching. The resulting deformation is a smooth interpolation between the mapping outside the mask, guided by the regularisation of the registration.

7.3.4 Quantitative analysis

The registrations were compared qualitatively, by visual inspection, and quantitatively by similarity of structures with the manually delineated gold-standard. The registration quality was assessed in both directions.

The structures were manually delineated on the pCT, MR_1 and MR_2 per patient by the same expert radiation oncologist, Dr Dhanasekaran Kittappa, Dr Swee-Ling Wong, and Dr Ruheena Mendes. A total of 12 structures were delineated on the CT, MR_1 and MR_2. This structure set provided an indication of how well the registration accounted for anatomical differences and positioning errors (Figure 7.4). This set consisted of vertebrae C3 and C5, mandible, thyroid cartilage (bony anatomy), spinal canal, brainstem, parotids (OARs), submandibular gland and sternocleidomastoid muscles (soft tissues).

The registrations were assessed in terms of the following using the following quantitaties: DSC, OI, FN, FP, DT and CoM (section 2.5.1.3 and 3.2.3.1); HE, properties of the determinant of the Jacobian [det(Jac)], and ICE.

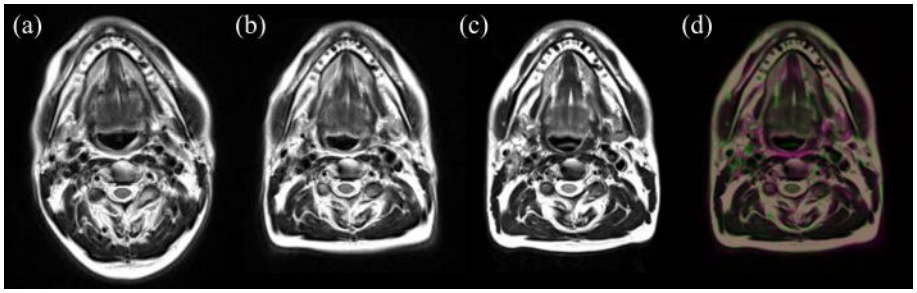

Figure 7.5: *Example of registrations: (a) MR_2, (b) MR_1, (c) MR_2 deformed to MR_1, and (d) overlay between MR_1 (magenta) and deformed MR_2 (green).*

7.4 Results

Analysing the quantitative results obtained for all the registrations performed, the combination of parameters that worked the best in the available datasets were BE=1% and CPS=12 mm for CT-MR_1, and BE=0.1% and CPS=12 mm for MR_1-MR_2. In CT-MR_1 registrations, it was preferable to use a higher weight of the BE than for MR_1-MR_2. This reduced the risk of the registration causing additional uncertainties in comparison with rigid-only alignment (such as deformation of bones). However, if the immobilisation is not present, those constrains should be relaxed to give the algorithm enough freedom to recover larger deformations. Since multimodal registrations had to capture larger anatomical changes the constrains had to be relaxed (Figure 7.5), and the properties of the DVFs reflect also this. A higher CPS in general resulted in DVFs with more desirable properties, which did not compromise the similarity between structures. For this combination of registration parameters, the results obtained for the quantitative evaluation of the DIR can be found in Table 7.1.

Additionally to the global results provided in Table 7.1, the results were also grouped and analysed by structure type. In MR_1-MR_2 registrations, the DSC values were 0.62±0.12, 0.77±0.08 and 0.84±0.07 for bony anatomy, soft tissues and OAR in DIR cases, and 0.4±0.3, 0.63±0.18 and 0.65±0.19 when using a rigid-only registration. The use of a rigid-only transform in MR_1-MR_2 registrations was not adequate, as for some anatomical structures the overlap could be close to zero due to the large anatomical changes and differences in positioning between pre and pos-treatment scans. For CT-MR_1 registrations the DSC values were 0.57±0.18, 0.79±0.06 and 0.81±0.04 for bony anatomy, soft tissues and OAR when using DIR, and 0.57±0.18, 0.74±0.09 and 0.79±0.04 when using rigid-only registrations. The results obtained with DIR and rigid-only registrations were very similar, with DIR performing marginally better in the soft tissue regions.

Table 7.1: *Mean values ± standard deviation of dice similarity index (DSC), overlap index (OI), false negative (FN), false positives (FN), distance transform (DT), centroid position error (CoM), harmonic energy (HE), properties of the determinant of the Jacobian [det(Jac)], and inverse consistency error (ICE) for CT-MR_1 and MR_1-MR_2 registrations. The results are averaged for all patients, structures or DVFs, and registration directions.*

	CT-MR_1	MR_1-MR_2
Geometric matching		
DSC	0.72±0.16	0.74±0.13
OI	0.73±0.19	0.75±1.5
FP	0.3±0.3	0.3±0.2
FN	0.27±0.19	0.24±0.15
DT_{mean} (mm)	-0.1±1.8	0.1±1.7
DT_{std} (mm)	2.2±1.2	2.3±1.6
$\|DT\|_{mean}$ (mm)	1.8±1.0	1.6±1.1
$\|DT\|_{std}$ (mm)	1.9±1.3	2.1±1.7
$\|DT\|_{95\%}$ (mm)	6±4	6±5
$\|DT\|_{max}$ (mm)	10±6	10±6
CoM (mm)	3±2	3±2
Characteristics of the deformation fields		
HE	0.15±0.01	0.39±0.06
$det(Jac)_{1\%}$	0.72±0.10	0.5±0.3
$det(Jac)_{99\%}$	1.32±0.09	1.7±0.3
ICE_{mean} (mm)	0.08±0.10	0.8±0.7
ICE_{std} (mm)	0.2±0.2	1.9±1.5
$ICE_{99\%}$ (mm)	0.9±1.2	10±8

7.5 Discussion

Promising results were found for multimodal and multitemporal registrations. For the soft tissues and OARs, values found were in agreement with results from other multimodal studies [213, 214], and comparable to monomodal or quasi-monomodal (CT-CT/CT-CBCT) studies [87, 221]. In spite of the large deformations between pre- and pos-treatment images, it was possible to achieve similar registration accuracy as for CT and MR in treatment position.

The registration of bony anatomy was poorer than for other types of structures. On one side, the reduced contrast between soft tissue and bone in MR difficulties the delineation of bones, particularly for complexly shaped structures such as the vertebrae, resulting in a non-ideal gold-standard. This low contrast also affects the quality of the registrations. However, the main interest in using MR is not to provide additional information on the bone anatomy (where CT is more adequate), but rather on the soft tissues. Thus misregistrations of the bones is of reduced importance when considering clinical applications and, in fact, in regions of higher clinical relevance, such as OAR and soft tissues, DIR performed in a higher level of accuracy. Nevertheless, the poor registration of the bones may affect nearby soft tissues so it is of importance to develop DIR strategies that account for the rigid behaviour of bony anatomy.

CT-MR_1 DIR slightly improved the anatomical matching in comparison to a rigid registration; however, the difference was not clinically significant. One must carefully tune its DIR registration to avoid introducing errors in this process. Further studies with a larger patient dataset are necessary to fully understand this additional uncertainty, and also to validate CT-MR registrations for patients that can not acquire MR_1 in treatment position (i.e., with considerable setup variation between scans).

The FoV clipping was found to limit the quality of the registrations. A clipped body contour reduces the ability to capture global deformations and generates unrealistic deformations within the patient near the edges of the FoV. For example, this was found to interfere in the registration of the mandible, which was very often clipped in the MR scans.

7.6 Current status and future work

The work here presented was a first step toward incorporating additional imaging into the radiotherapy pathway; thus some aspects of the current work were not finished. Two particular aspects were investigated but not included in this chapter:

1. Only work on registering CT-MR_1 with similar positioning was reported in this thesis. However, a similar optimisation was performed using two datasets where the MR_1 was not acquired in treatment position. These datasets served as training

sets for the optimization of registration strategies, which differed from the cases where the patient could be immobilised in the MRI scanner. The most important conclusion of this study was that the constrains of the registration had to be relaxed to be able to account for changes in position; however, no quantitative data was available of this evaluation due to lack of complete delineations from physicians.

2. Unlike the CT-to-CBCT registrations used in the previous chapters, in CT and MR it was possible to unequivocally identify points and OAR relevant to HN treatments. The target registration error (TRE) was measured, and analysed using the mean and root-mean square error of the absolute TRE. The components in the right-left (RL), anterior-posterior (AP) and superior-inferior (SI) were also assessed separately. These results were also not reported as complete identification of landmarks was not concluded by the completion of this thesis. Nevertheless, all the analysis code was produced and tested.

Traditionally, MRI was used to acquire additional anatomical information, since unlike CT it provides high contrast in soft tissues. Nowadays MRI technology is rapidly evolving from anatomical and structural to dynamic, functional and metabolic imaging. Functional MRI sequences include DCE-MRI, DW-MRI and MRS [199]. The work here presented focused on anatomical information only, so future work will also focus on tracking functional information associated with the anatomical scans, with a particular interest in DW-MRI. DW-MRI measures the diffusion of water in tissue. Water motion is not random in tissue, but instead modified by flows within conducts and interactions within the cells and with the extracellular matrix [222]. Movement of the tissue water molecules between two gradients results in dephasing, depicted as signal loss. This signal loss will be proportional to the amount of water molecule movement and the strength of the gradients. By repeating the sequence with different gradients, one can quantify the observed signal loss using the ADC [201]. There appears to be a correlation between the ADC values, tumour cellularity and tumour grade. ADC values can be measured before and during treatment to demonstrate the presence or absence of therapy-related changes in tumour tissue architecture [199].

Using DIR to map functional information is still a unexplored topic in the literature. Galban et al. evaluated mid-treatment MR as an early biomarker for outcome, using a semi-automatic algorithm to propagate contours between time points [223]. A step forward from this study would be to look at sub-regions of the tumour, which would also require an accurate one-to-one mapping. Functional acquisitions are acquired co-registered to anatomical acquisitions, but further corrections may be necessary to deal with artefacts and patient motion. In the same study functional-to-anatomical DIR is performed to deal with susceptibility artefacts and patient motion. However, it is probably more adequate to first correct the functional images using specialised methods for that purpose as DIR can introduce additional uncertainties [224, 225].

One of our main interests is to use multimodal and multiparametric MR imaging in

the context of ART applications, which is becoming an increasingly relevant topic with the advent of the MRI-LINAC [226]. The present study is not ideal to validate the use of DIR for such applications, as the multiple MRs were not acquired throughout the treatment. Such a dataset was not available at UCLH or easily available from other collaborators. However using a pos-treatment MR results in more challenging registrations, and therefore MR_2 can be considered a surrogate for MR acquired during treatment. This is however only true when considering the mapping of healthy tissues. Tracking of tumour volumes has to still be properly validated when the MRs are acquired throughout the course of radiotherapy. However, this is not a trivial matter due to the complexity of the mechanisms of tumour microscopic response. The challenges in tracking and validating tumour progression are the motivation behind the work presented in chapter 8, where the concept of using an *in vitro* cancer model as a controllable and deformable bio-phantom is explored.

CT-to-MR and MR-to-MR DIR are of interest in radiotherapy for other applications, such as MR-based treatment planning and estimation of attenuation corrections for PET/MR scans. MRI-LINAC systems are becoming available in clinical settings, and synthesising CT from MR is one of the potential applications of CT-to-MR and MR-to-MR [227]. Quantitative PET reconstruction requires correction for photon attenuations using an attenuation coefficient map, that is a measure of the electron density (i.e., CT information) [228].

7.7 Conclusion

In this chapter the use of an open-source DIR algorithm was investigated for the registration of CT and MR datasets from HN patients. The results founds were preliminary but promising, which allowed to identify the limitations of current DIR algorithms and current protocols of MRI acquisition. This was a first step toward incorporating additional imaging into the radiotherapy pathway.

Chapter 8

A novel artificial cancer mass model for imaging applications

> To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science.
>
> Albert Einstein

In this cancer models were investigated as a bio-phantom for multimodal imaging applications. Tissue engineering techniques were explored to produce living samples MRI-friendly; the design process of the samples and the initial imaging and characterisation are the focus of this chapter. This chapter, unlike the previous ones, has a more descriptive character. It describes in detail the exploratory work performed in this topic to guide future studies.

The work in this chapter resulted in the following output:

- C. Veiga, T. Long, B. Siow, M. Loizidou, G. Royle, and K. Ricketts, "MO-F-CAMPUS-I-04: Magnetic resonance imaging of an *in vitro* 3D tumor model," Med. Phys. 42(6):3579 (2015).

8.1 Introduction to tissue engineering

One of the major obstacles in translating multimodal and functional imaging to routine clinical practice is the lack of clinical evidence of the benefits of doing so. Even though numerous benefits of using additional multimodal and multiparametric imaging (particularly MRI) in radiotherapy have been reported by several research groups, ranging from

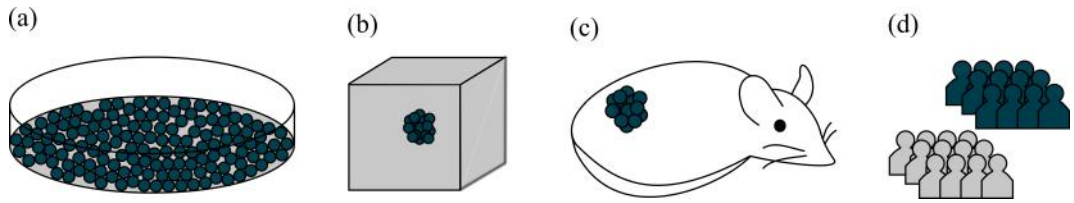

Figure 8.1: *Platforms to study cancer and therapies: (a) monolayer cells lines, (b) in vitro 3D cancer models, (c) xenograft and, (d) clinical trials.*

animal ([229–232]) to patient studies ([59, 210, 223, 233]), there has yet to be a translation from research to routine clinical practice. Additional imaging is costly and poses additional risks to the patients, so further evidence of the benefits of acquiring it is necessary. There are several platforms to study cancer and therapies (Figure 8.1). Ground-truth evidence is gathered in clinical trials, which are performed on a small sample of patients and require ethics approval. Most of our current knowledge of cancer biology and the effect of therapeutic treatments outside of clinical trials comes from pre-clinical studies, from which the two most common models are *in vitro* monolayer cell lines (2D) and *in vivo* animal trials (xenograft). These models attempt to mimic the biological characteristics of the tumour microenvironment outside the human body. The tumour microenvironment consists of the cancer cells, which interact dynamically with the surrounding stroma. The stroma contains non-cancer cells, secreted soluble factors (such as growth factors) and non-cellular solid material that provides structural support and biomechanical properties, the ECM. The ECM of connective tissues represents a complex combination of diverse protein families: collagen, fibronectin, elastin and laminin.

Traditional *in vitro* systems involve growing cancer cell lines in monolayers in cell culture (2D). The main advantage of this simple model is to be highly reproducible and responsive to drugs and radiation [234]. 2D studies contributed greatly to our current understanding of carcinogenesis and tumour response to treatments [235]. However, the non-existence of their native microenvironment often does not accurately forecast a tumour's *in vivo* response [236]. This approach lacks essential cancer characteristics: complex interactions between cancer cells and stromal cells or ECM components, hypoxia and angiogenesis drivers and architectural characteristics [235].

The most common *in vivo* model is the xenograft, which consists of injecting human cancer cells or small fragments from cancer specimens in the laboratory mice (*Mus musculus*), and allowing the tumour to grow. A great advantage of this method is that both cancer tissue and surrounding stroma can be transplanted, mimicking the complexity of the human tumour environment. The main disadvantages of this method are its laborious nature, the need for animal facilities, ethical approval, and the resulting poor predictions [234, 237].

As an alternative to animal and patient data, the use of novel artificial cancer masses to provide evidence of the benefits of additional imaging during radiotherapy is suggested

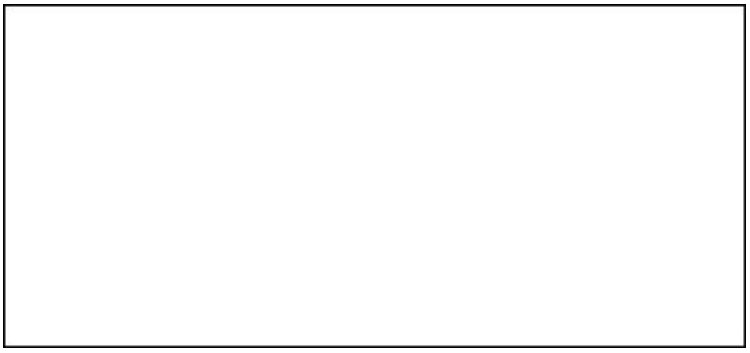

Figure 8.2: *The three elements of tissue engineering: cells, matrix (or scaffold) and biochemical and physical signals. Adapted from [242].*

and investigated in this chapter. *In vitro* tumour models are invaluable systems for studying the dynamic and progressive behaviour of cancer under controlled conditions [238]. 3D *in vitro* models have been used in cancer research as a compromise between the simplicity of 2D cultures and the complexity of mice models. Intermediate 3D models were firstly developed, known as multicellular tumour spheroids [239] and gel embedding [240], and are able to mimic limited aspects of the tumour biology. Such models have an intrinsic limitation in size due to the lack of mechanisms of oxygen transport. Recently tissue engineered 3D complex tumour models have been built to be biomimetic (both cell-cell and cell-ECM interactions are reproduced), aiming to accurately replicating the native *in vivo* scenario in which they are found [241]. Cells cultured in 3D configurations differ greatly from cells cultured in 2D monolayers in terms of cell density, ECM synthesis, cell surface receptor expression, cell contraction, intracellular signalling and metabolic functions [238]. Three key elements define a tissue engineered structure: the properties of the cells cultured, the surrounding matrix (that mimics the ECM), and their biochemical and physical properties (Figure 8.2).

Previous work by the Division of Surgery and Interventional Science (UCL) enabled the construction of 3D samples that closely mimic colorectal cancer [243]. In chapters 3 and 7 the ability of using DIR to map contours between scans was evaluated, but only for OARs. Propagation of target volumes is more challenging due to the complexity of the tumour microscopic structure and of the biological responses that occur as effect of treatment. A quantifiable and deformable bio-phantom can potentially provide some insight into this clinical problem. Thus the motivation behind the establishment of a MRI-friendly model in the context of this thesis was to create a controllable gold-standard for validating DIR of the tumour progression during radiotherapy treatments. Thus, this chapter focuses on exploring the concept of an engineered tumour model compatible with multimodal and sequential imaging studies. This work is inserted in a larger research interest in the application of the tumoroid model. The main focus of this chapter was MR imaging; other research interests of the tumoroid research group include treatment response, dose enhancement studies, and other imaging techniques. My contribution to the whole project was the theoretical design of the samples, planning of the experimental

pre-clinical MRI setup and, initial characterisation of the samples in terms of MRI contrast. Some preliminary work was also performed on CT imaging (section 8.4.3.3), but this was not prioritised due to the poor tumour contrast characteristic of CT imaging in comparison with MRI.

The ACM model has the potential to act as a challenging 3D imaging bio-phantom:

(i) The model has shown characteristics similar to those found *in vivo* tumours.

(ii) The samples are easy to produce, and can be used in imaging and/or treatment studies without the need of ethics approval.

(iii) The biological characteristics of the tumoroids are controllable and reproducible, making them a perfect test subject for repetitive and sequential studies.

The use of 3D tumour models as bio-phantom for radiotherapy and medical imaging applications has so far been poorly explored. Few studies available focus on routine and clinical imaging modalities using less complex 3D bio-phantom as tool to verify, for example, nanoparticles as MRI contrast agents [244]. Furthermore, in previous work conducted within our group gold-nanoparticle concentration was measured on 3D ACMs using non-destructive X-ray fluorescence technique [245]. The imaging-friendly artificial cancers can potentially be used for a multitude of high impact applications in radiotherapy:

- To calibrate imaging acquisitions of different imaging systems, which could be of commercial interest.
- To provide pre-clinical evidence of the benefits of additional imaging and different treatment choices, via multimodal and sequential imaging and treatment, and correlation with histology studies.
- To further individualise radiotherapy, by deciding on the treatment approach after testing the efficiency of different treatment modalities in patient-specific artificial cancer masses.

8.2 Engineering of a tridimensional cancer model

The ACM model developed at the Division of Surgery and Interventional Science (UCL) consisted of cancer cells seeded in a collagen matrix, which were compressed to increase the density, and embedded in an uncompressed collagen gel that can contain sparse fibroblasts and healthy cells. This model had shown a number of characteristics similar to *in vivo* cancer masses, such as migration, hypoxia, and release of pro-angiogenic factors capable of initiating cancer-related angiogenesis [235].

In this section the standard protocol to produce the 3D artificial cancer mass is described as an introduction to tissue engineering. It consists of the two main steps: cell

culture and collagen gel preparation.

8.2.1 Cell culture

The cell line used was the human colon adenocarcinoma cell line (HT29). The cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) with 1000 mg glucose/L, L-glutamine, NaHCO and pyridoxine HCl (Sigma-Aldrich, St Louis, MO, USA) supplemented with 10% foetal bovine serum (FBS) (First Link UK Ltd, Birmingham, UK) and 1% penicillin/streptomycin (P/S) (GIBCO, Invitrogen, Paisley, UK). Cells were maintained under sterile conditions as 2D monolayers in 100% humidity, 5% CO_2 in air at 37°C in an incubator.

The medium maintains cells in tissue culture by providing the nutrients the cells require to stay alive and healthy (amino acids, salts, glucose and vitamins). It must be changed frequently (at least twice a week) to replace the nutrient levels and remove cell waste. A standard cell maintenance protocol was followed (Appendix B).

Once the cells are seeded into a flask, they adhere on the bottom surface. The number of cells increases as the cells undergo mitosis to populate the available space. The term "confluency" is used to numerically describe this process: x% confluency describes the situation when x% of the flask surface is covered in cells. Once 100% confluency is reached the growth eventually stops due to contact inhibition, lower availability of nutrients and excessive waste. To maintain a healthy cell culture the cells are subcultured before reaching 100% confluency (between 85-95%). This process is called passaging, which involves enzymatically detaching the cells from the flask surface, and transferring a fraction of the cells to new flask(s). A standard cell passaging protocol was implemented throughout (Appendix C).

An accurate number of cells embedded in the collagen matrix are necessary to assure reproducibility in manufacturing the bio-phantoms. The number of cells in a cultured flask can be measured using a haemocytometer and a standard cell counting protocol. The total number of cells is then re-suspended in the necessary volume of DMEM to quantify the cell concentration when seeding the collagen matrix (Appendix D). A haemocytometer consists of a counting chamber in a thick glass slide covered with a thin glass coverslip. The chamber is laser-engraved with a grid of perpendicular lines visible under microscope. The volume of each grid unit is known, so once the chamber is filled with cells in suspension it is possible to count the number of cells per volume unit. The concentration of the cells should neither be too high or too low, as high concentrations result in cell overlapping and low concentrations in higher statistical errors. The accuracy of this manual counting technique ranges from 7.1 to 15.6% [246].

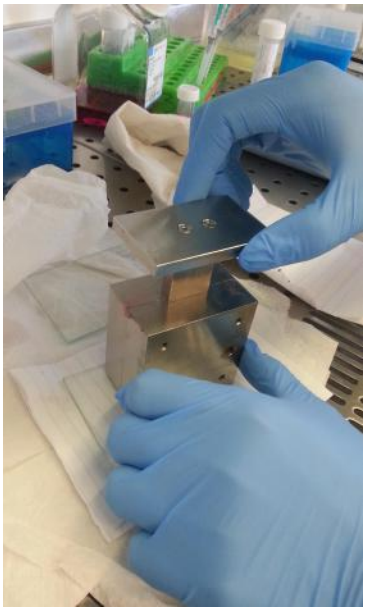

Figure 8.3: *Photo of the mould and plunger system for plastic compression. The mould is placed on top of the nylon and steel meshes.*

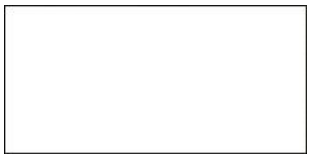

Figure 8.4: *Photo of the artificial cancer mass after plastic compression, and its dimensions. Courtesy of Tarig Magdeldin.*

8.2.2 Collagen matrix

The ACMs were manufactured following a standard protocol to seed the cancer cells in collagen hydrogel and applying plastic compression to increase the density (Appendix E). Collagen hydrogel (rat tail collagen type I, 2.04 mg/ml in 0.6% acetic acid, First Link UK, Birmingham, UK) is mixed with 10× concentrated minimum essential medium (MEM) (Invitrogen, Waltham, MA, USA). The solution is neutralised in a drop-wise manner with sodium hydroxide (NaOH). The solution changes from yellow to bright pink when the optimum pH level of 7.3 is reached. The cell suspension is then added to the solution. In the original ACM protocol the collagen is populated with 6.4×10^6 cells. The final solution has 4 mL, composed of 3.2 mL of collagen, 0.4 mL of MEM and 0.4 mL of cell suspension.

The thoroughly mixed solution is transferred to a mould and left to incubate at room temperature for 30 minutes, allowing for gel setting. Once set, the ACM density is increased by plastic compression (using a mould and plunger as shown in Figure 8.3) and self-compression. The plastic compression process is repeated from both sides of the ACM to homogenise the compression within the volume.

The compressed ACMs can then be cut using a sterile surgical blade (Figure 8.4). The

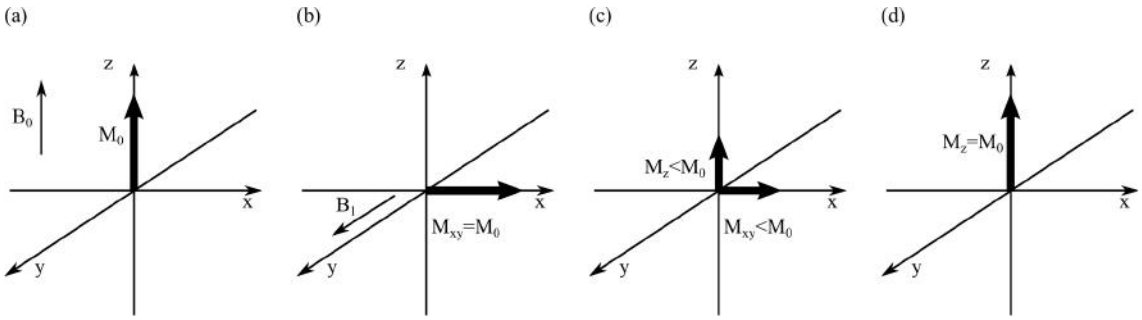

Figure 8.5: *MR relaxation to equilibrium. (a) At equilibrium the magnetisation (M_0) is aligned with the external magnetic field (B_0). (b) A 90° pulse (B_1) is applied for enough time to rotate the M_0 into the x-y plane. (c) With time the magnetisation relaxes to equilibrium. (d) Some time later ($t>>T_1$) the magnetisation has returned to equilibrium state.*

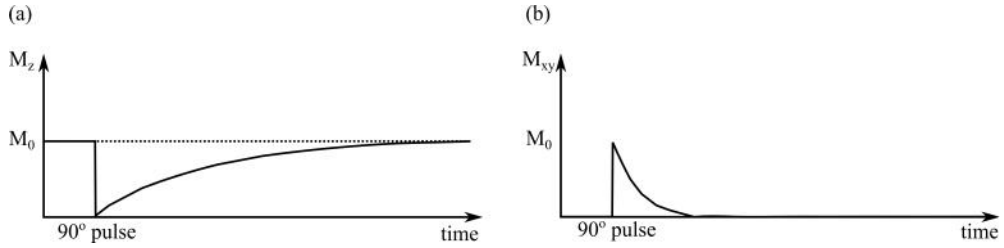

Figure 8.6: *Components of the magnetisation in (a) z axis and (b) x-y plane before, during and after the 90° RF pulse.*

pieces are then embedded in uncompressed collagen, which can be populated with other cells. The samples are then covered in DMEM and incubated. The medium is removed and replaced regularly.

8.3 Physics of magnetic resonance imaging

8.3.1 Contrast mechanisms of conventional magnetic resonance imaging

MRI uses the natural properties of the hydrogen atoms, a major constituent of the human body to generate images of high soft tissue contrast. When a sample is inside the magnet it acquires a very small magnetic field ("magnetisation", M_0), which aligns with the strong field of the scanner (B_0). Radio-frequency (RF) pulses are used to change the direction of this magnetisation to the x-y plane, described by the flip angle (α). Following the RF pulse the system relaxes back to equilibrium: dephase in the x-y plane (signal decay) and the macroscopic magnetisation realigns with the field [247–249]. Figure 8.5 describes schematically the process of relaxation to equilibrium, which occurs via two processes at different rates, spin-lattice (T_1) and spin-spin (T_2) relaxation times. The components of the magnetisation can be represented graphically (Figure 8.6).

1. T_1 relaxation: after the RF pulse, the magnetisation gradually realigns with the external magnetic field in a time characterised as spin-lattice relaxation time. The

hydrogen nuclei lose their magnetic energy to the surroundings (the lattice). The equation for the recovery of M_z following a 90^opulse is:

$$M_z(t) = M_0(1 - e^{-t/T_1}) \tag{8.1}$$

2. T_2 relaxation: MR signal that is in the x-y plane (M_{xy}) evolves by decaying away at an exponential rate defined by the spin-spin relaxation time T_2. The rate at which M_{xy} is lost is given by:

$$M_{xy} = M_0 e^{-t/T_2} \tag{8.2}$$

T_2 decay is generally shorter than T_1 as it governed by more than one effect: (i) exchange of energy between nearby spins and (ii) presence of nearby magnetic molecules that perturb and change locally the magnetic field. In real systems the magnetic field is never perfectly uniform, and the existent inhomogeneities also accelerate the decay of M_{xy}. The combined T_2 and magnetic field inhomogeneity is known as T_2^*.

T_1 and T_2 are characteristic of the tissues, but are not constant as they vary with the strength of the external magnetic field and temperature. For water at 1T both have a value of 2500 ms [250].

The magnetisation is measured in the x-y plane by detecting the voltage it induces in a receiver coil. This signal is known as free induction decay (FID) and it decays exponentially to zero since the protons are de-phasing and relaxing back to the equilibrium. In practical terms the FID is not measured, but rather echoes created from the signal.

8.3.2 Pulse sequences

MR images are produced using a pulse sequence, which is stored in the scanner computer. The sequence contains the RF pulses and gradient pulses which have carefully controlled durations and timings. There are two types of pulse sequences: spin-echo (SE) and gradient echo (GE). The choice of repetition time (TR) and echo time (TE) define the type of contrast (Table 8.1). Other parameters such as inversion time (TI) and flip angle (α) are also used to define particular sequences.

SE uses two RF pulses, usually 90° and 180° pulses, to create an echo which measures the signal intensity. In such sequences the accelerated T_2^* is reversible by the application of the 180° pulse, that refocuses the spins to form a SE. SE produces the images with higher quality at the cost of longer acquisition times.

GE uses a single RF pulse followed by a field gradient reversal to generate an echo without an 180° pulse. Like SE sequences, any type of contrast can be generated (Table 8.1), but these are much faster acquisitions. Thus, they are more influenced by field

Table 8.1: *Choice of repetition time (TR), echo time (TE) and flip angle (α) and contrast generated for spin-echo (SE) and gradient echo (GE) acquisitions. For GE the TR is always short (<750 ms).*

SE				GE	
	TE			TE	
TR	Short (<40 ms)	Long (>75 ms)	α	Short (<15 ms)	Long (>30 ms)
Short (< 750 ms)	T_1	N/A	Small (<40°)	PD	T_2*
Long (> 1500 ms)	PD	T_2	Large (>50°)	T_1	N/A

inhomogeneities and timing of the parameters such that the amplitude of the gradient echo is determined by T_2^* decay.

Within the two major families of pulse sequences (SE and GE) there are some well-known acquisitions. Relevant to this project are the fast low angle shot (FLASH) and rapid acquisition refocused echoes (RARE). The first is a fast GE with low flip angles, and the second refers to multiple fast SE.

8.3.3 Measurement of T_1 and T_2 relaxation times

T_1 and T_2 are properties of the tissue, and can be measured using imaging. One method of T_1 measurement requires a series of inversion-recovery (IR) sequences with varying TI. IR starts with a 180° inversion pulse that inverts M_0, which then starts to recover. A time is waited after the 180° pulse (known as TI) before continuing with the SE imaging sequence. When the 90° pulse is applied some of the signals may still be negative, and a mixture of positive and negative echoes will be formed. TR must be at least five times the longest T_1 to allow full relaxation between inversion pulses. When this condition is met, the signal given T_1 is extracted by fitting the data to the curve:

$$M_z = M_0(1 - 2e^{-TI/T_1}) \tag{8.3}$$

T_2 is measured by using a train of SE sequences to measure the signal intensity at varying TE. T_2 is extracted by fitting the data to the curve:

$$M_{xy} = M_0 e^{-TE/T_2} \tag{8.4}$$

8.4 Design of an artificial cancer mass for magnetic resonance imaging

8.4.1 Design specifications

Ideal ACM samples should provide imaging contrast both in anatomic and functional sequences; further engineering is necessary to achieve this goal. The specifications desired for the engineered samples are:

(i) The tumoroids should provide image contrast in different imaging modalities, such as CT and multiparametric MR.

(ii) The contrast generated should be quantifiable based on its biological properties (such as cell density, cell viability, etc).

(iii) The tumoroids should be compatible with sequential imaging (i.e, multitemporal).

(iv) Changes in the biological properties of the tumoroids (for example, due to treatment) should be within the sensitivity of the imaging system.

8.4.2 Biological properties of the samples

8.4.2.1 Cell density

The major limitation of the original model by Nyga et al. was the low cell density of the ACM, which reduced the generation of MR contrast. The cell density can be increased by varying three different parameters during the ACM production: cell density, collagen density and incubation time.

Cell seeding Increasing the number of cancer cells seeded in the collagen matrix is the most straightforward method to increase the cell density. However, empirically it is found that when more than 30-50$\times 10^6$ cells were seeded the collagen gel no longer sets properly. This could be explained due to the large number of cells causing interference during the collagen fibrillogenesis, and/or non-optimal pH conditions when the high concentration of cells is added. Additionally, previous studies show that cells growth rate in the ACM from day 7 to 14 was independent of the initial seeding density (Figure 8.7). For imaging applications it is desirable for the cell number to be as high and as stable as possible during the first days of the tumoroid life.

Collagen density The collagen density is increased by plastic compression: as the water is squeezed out by external pressure, the volume of the construct is reduced which effectively increases the cell density. The weight and duration of the plastic compression

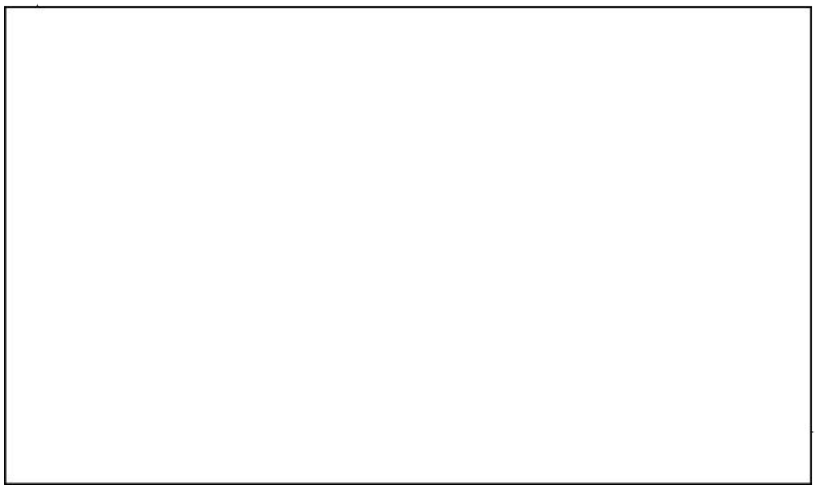

Figure 8.7: *Cell densities of HT29 measured over 14 days in a partially compressed ACM for different cell seeding values on each time points. Adapted from [235].*

step must be set to maximise the final density of the samples. Brow et al. had found that significantly extending the plastic compression time did not increase fluid loss but does reduce the inter-construct variance [251]. Additional weight, however, did increase the collagen density. Therefore several methods to optimise the plastic compression process were investigated, which were tested in the laboratory by Tong Long (Division of Surgery and Interventional Sciences, UCL):

1. Use the original mould (Figure 8.3) and increase the compression time. Parallel experiments to measure the collagen density as a function of the compression time were performed, and results similar to Brow et al. were found [251].

2. Design a new mould. Varying the surface could affect the expelling of water and allow for additional compression. A mould with varying surface was designed, and smaller surfaces resulted in less compressed constructs. Therefore, the initial mould was used in the imaging experiments. Additional information on work on this point can be found in appendix F.

3. Use the available mould and increase the load. The process had to be done in steps of increasingly external pressure or the construct would loose its integrity. This allowed to increase the collagen density, defined as the ratio between dry and wet weight, from 20±3% to 40±2% (Figure 8.8).

The technique chosen to maximise the plastic compression of the samples imaged used the previously available mould with additional load. The main limitation of this approach (which will be discussed further on section 8.5.2.2) was the poor reproducibility achieved with the technique. Following preliminary studies of the samples, a standardised protocol was followed to attempt to produce reproducible samples. Figure 8.9 presents the schematic diagram of the protocol followed to ensure higher reproducibility between samples. The plastic compression maximum height is controlled via a construct of de-

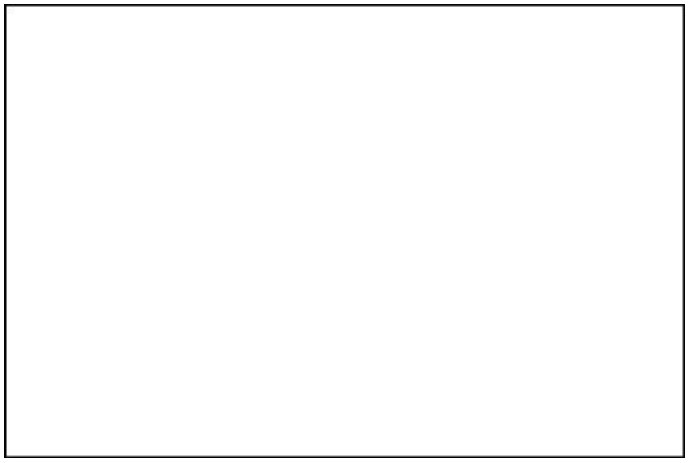

Figure 8.8: *Collagen density (ratio between dry and wet weight) for different compression times and additional force (C). Courtesy of Tong Long.*

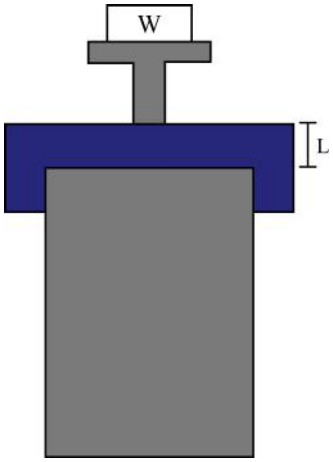

Figure 8.9: *Plastic compression with additional constructs for higher reproducibility of the collagen density: construct of fixed dimension (L) to constrain the compression and additional weight (W).*

fined length (L) that stops the compression after that threshold is reached. This constructs can be 3D printed to customise the size and ensure adequate fitting to the mould. The size varies between sides being compressed, being longer on the first side ($L_1 > L_2$) . A fixed weight (W) can be added to the top of the plunger to generate the additional compression. The dimensions and weights were not fully optimised; values of L_1=6 mm, L_2=4 mm were chosen empirically, and the additional weight was applied manually.

Incubation time Once the gel is set, it should consist of a uniform distribution of cancer cells in a collagen gel. As the ACM is given time to stabilise, the cancer cells keep proliferating and start to migrate from the centre, which is hypoxic, forming at the surface high density clusters. With time, the cells start to detach from the dense collagen and invade the surrounding stroma (Figure 8.10). This is similar to *in vivo* behaviour of cancer invasion. Given time, the cells will contract the collagen gel, eliminate fluid and improve the mechanical properties [252].

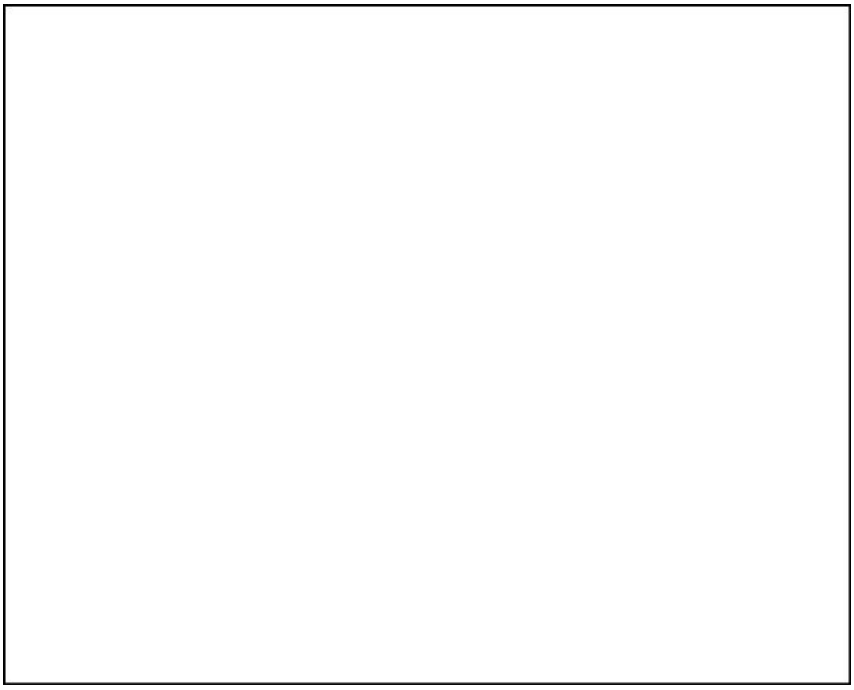

Figure 8.10: *Days (a) 1 (b) 7 (c) 14 and (d) 21 (using at seeding $2{\times}10^6$ cells per ml of gel). The microscopy images of the sections were taken at 40× magnification. Courtesy of Tarig Magdeldin.*

The incubation time should be appropriate for the gel to stabilise; short enough to reduce the risk of contamination and adequate considering cell proliferation. As seen in Figure 8.7, tumoroids with very high cell number at seeding become less dense with time. Due to the high number of cells of our constructs, the tumoroids were incubated for 24 hours before imaging. This allowed stability of the construct, higher cell density homogeneity and higher viability of the cells at the time of the imaging sessions.

8.4.2.2 Sample fixation

Fixation is a chemical process by which biological tissues are preserved from decay, thereby preventing autolysis or putrefaction. Fixation terminates any on-going biochemical reactions, and may also increase the mechanical strength or stability of the tissues. However, chemical fixation causes the MR properties of biological tissues to be different from those found *in vivo*. Moreover, these properties change as fixation time elapses. For T_2 there is a general reduction of signal with time ([253, 254].

Correlating images from fixed and living samples is challenging, and may be affect the ability to generate MR contrast. Fixation however facilitates the logistics of optimising samples and imaging protocols. The decision was to perform all the following imaging experiments with living samples since this was also a necessary condition for sequential imaging experiments.

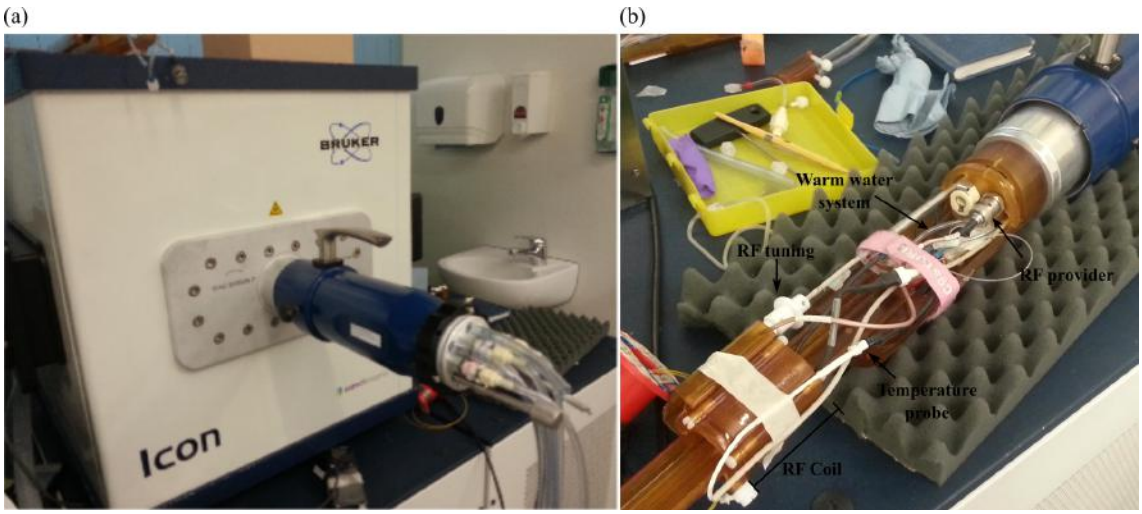

Figure 8.11: *(a) Bruker ICON^{TM} MRI scanner system and (b) composition of the coil-holder and its electronics.*

8.4.3 Design of the imaging experiments

8.4.3.1 Magnetic resonance system specifications

The images were acquired at CABI using the ICON^{TM} (Bruker Corporation, Billerica, MA, USA) MRI scanner. It is a compact small footprint high-performance MRI system for pre-clinical research and *in vivo* imaging (Figure 8.11a). The ICON^{TM} operates at a 1 T field strength.

The magnet bore is horizontal and Figure 8.11b shows the system where the RF coil is embedded. The piece is equipped with a system of flowing water that allows to maintain the temperature of the samples. The temperature of the water can be selected, measured and monitored in real-time. This system was set so that the tumoroid is at an ideal temperature as if inside the incubator (T≈32^{o}C).

8.4.3.2 Experimental setup and sample holder

An optimised setup for MRI takes in consideration two aspects: (i) the size to the coil should be as small as possible while enclosing the sample and (ii) the filling of the coil should be maximised (i.e., minimise empty volumes). The tumoroid samples were cultured in a 7 mL bijou tube (L=38 mm, D=18 mm) that fits inside the smallest RF coil available (L=40 mm, D=23 mm, as Figure 8.11). The bijou tube was chosen due to its adequate dimensions, MR-compatible materials and sterile availability.

Since the sample does not fill the whole coil, a sample holder was designed to maximise the filling, as shown in Figure 8.12 and created using a 3D printer. A small aperture at the top is included to facilitate the contact between the temperature probe and the sample (alternatively a non-sticky conductive material could be used as interface between the

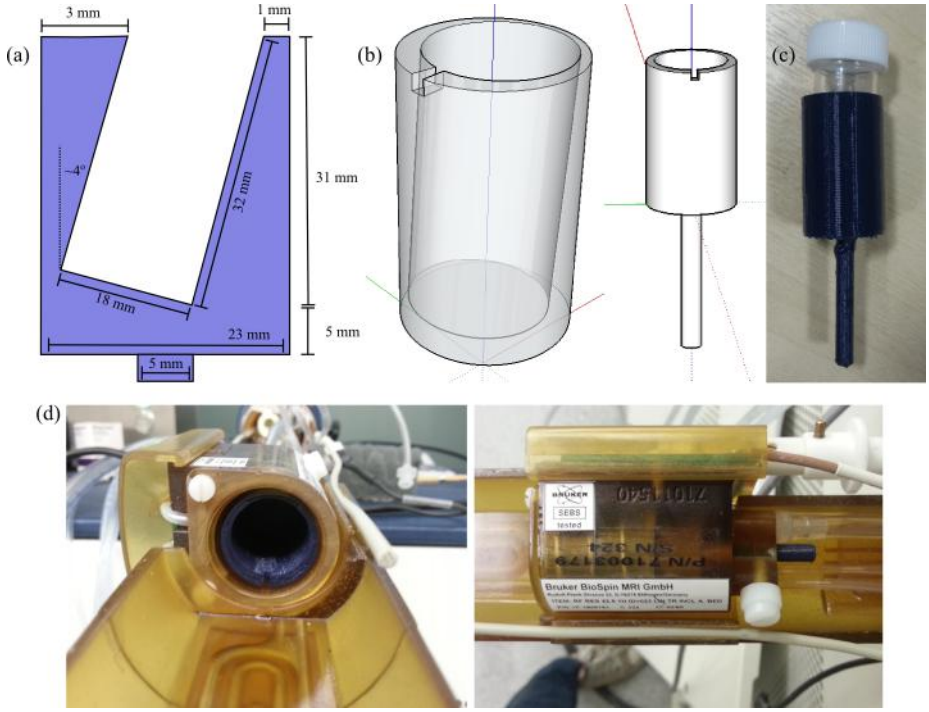

Figure 8.12: *Sample holder for MR imaging (a) schematic diagram, (b) 3D sketch and, 3D printed sample holder inside the holder: (c) frontal and (d) superior views.*

tube and probe). Additionally, it provides a guideline for the orientation of the holder inside the coil. An additional stick exits the bottom of the holder, which can be inserted inside an available coil hole and screwed for stability, allowing to switch between samples during the same experiment without moving the holder. The design was iterated a couple times to ensure appropriate dimensions after 3D printing. Dr. Robert Moss and George Randall (Department of Medical Physics & Biomedical Engineering, UCL) provided 3D printing and workshop expertise to finalise the prototypes of the sample holder.

Within the bijou tube, the filling must also be maximised which is done by overfilling the tube with DMEM. Since the magnet bore is horizontal, the tube must be in horizontal position as well. Because it is impossible to fully fill the bijou tube (there is always a residual air gap), a small air bubble is generated inside the tube. When the tube is in horizontal position this bubble can come in contact with the collagen. The collagen matrix starts to shrink and loose its integrity when in contact with air. To avoid this the sample holder hole can be tilted by a small angle to minimise the risk of bubbles coming in contact with the gel (Figure 8.12).

8.4.3.3 Fiducial markers

Fiducial markers within the samples are necessary to guide image acquisition and are of particular interest for imaging registration studies as a reference point. An ideal marker would be irregularly shaped, be positionable such that it clearly identifies where the ACM is located within the stroma without affecting the tumoroid morphology and biological

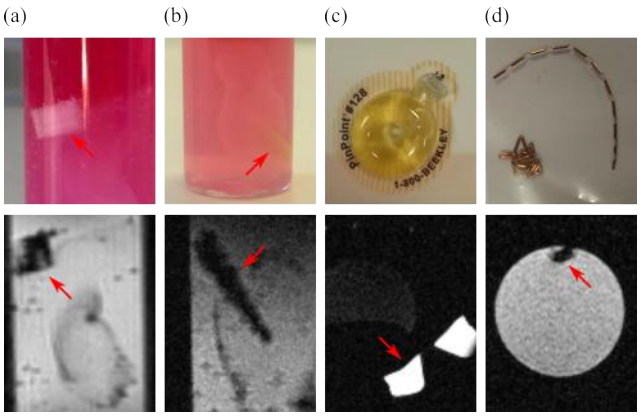

Figure 8.13: *Magnetic resonance imaging compatible markers tested: (a) plastic, (b) wood, (c) PinPoint® and (d) Gold Anchor™.*

properties (biocompatible), do not generate any MR signal and/or cause susceptibility artefacts. If internal, the markers have to be sterile or autoclavable. With the aim of making the tumoroids friendly for additional imaging modalities as a long-term goal, the markers would ideally be MR and CT compatible.

Several types of internal and external markers were investigated, both made in-house or commercially available (Figure 8.13a), simultaneously with the experiments that will be presented in section 8.5:

- Plastic marker. This marker was made in-house and irregularly shaped. It can be introduced to the ACM during the plastic compression or to the surrounding stroma. Polypropylene was used in the preliminary assessment, but a 3D printed marker made of polylactic acid (PLA) is also a comparable alternative.

- Wood marker. This marker was made in-house and had a pointed shape. It can be introduced to the ACM during the plastic compression or to the surrounding stroma.

- PinPoint® (Beekley Medical, Bristol, CT). This is a CT/MR image registration external marker developed commercially for clinical use. It contains a 1.27 mm diameter center hole that can be used as landmark for image registration. Due to its large dimensions and non sterile availability, it had to be attached to the bijou tube (i.e., externally to the sample).

- Gold Anchor™ (Naslund Medical, Huddinge, SW). This is a CT/MR image registration fiducial marker developed commercially for clinical use. It consists of a 0.28×20 mm rod that has two possible configurations (straight or ball shaped). It is normally introduced in patients using a sterilised needle. However, the Gold Anchor™ samples available were not sterile and were added to an acelular tumoroid during plastic compression.

Figure 8.13b shows MR images acquired of the markers. The plastic marker was

found to be the more adequate for MR-only applications as it was clearly identifiable and did not generate any visible artefacts in T_1, T_2 and T_2^* acquisitions. It can also have a shape fully personalised using a 3D printer. The wood marker had the disadvantage of absorbing water with time, and therefore it was found to also generate signal, which is sub-optimal. Its shape was however desirable, as it clearly points towards the ACM. The PinPoint® was found as non-usable, since not only its larger dimensions required a larger coil (and therefore the setup would be less optimal) but also it generated positive contrast that saturated the signal, removing the ability to distinguish details within the collagen samples. Finally, the Gold Anchor™ was used in a ball shaped arrangement, and even though it was easily identifiable it generated susceptibility artefacts. These artefacts were however not severe, and due to the high resolution of the system they could be reduced by using a single strand of the marker in straight configuration. For CT/MR applications this type of marker is the more desirable, and it was imaged using the X-TEK Real-time x-ray benchtop μCT (Nikon Metrology UK Ltd, Hertfordshire, UK) (Figure 8.14). The images were acquired with a resolution of 27×27×27 μm^3, and unlike the ICON™ the samples were placed vertically inside the scanner and therefore no sample-holder system was necessary for these preliminary acquisitions. Daniel O'Flynn (Department of Medical Physics & Biomedical Engineering, UCL) was responsible for acquiring the data on this system. The high attenuation of the markers made them more susceptible to artefacts during CT reconstruction (Figure 8.14c). However, that can also be minimised by using the marker in straight configuration. Its sterile availability is also an advantage of this markers, while its increased cost is a limitation. Simultaneously to the imaging of the marker, an acellular and 20×10^6 HT29 cell tumoroid were also imaged. These preliminary results had no evidence of x-ray contrast of the collagen gels, indicating the need of further engineering for CT imaging. In summary, a plastic marker is the most efficient for MRI, while the Gold Anchor™ is more versatile for multimodal applications. Further studies are necessary to optimise their shape and size, and assess their biocompatibility.

8.4.3.4 Sample transportation and storage

The samples were produced at the Division of Surgery and Interventional Sciences at the Royal Free Hospital and imaged at CABI, located within the UCL main campus. The samples had therefore to be transferred and stored for the imaging sessions. At CABI no incubator equipment was available, and therefore the samples were kept in an oven when not being imaged. This temperature had to be lower than a normal incubator (≈35°C) due to gradients of temperature inside the oven, particularly near its walls.

From the several imaging sessions performed, it was clear that the samples were very sensitive to high variations in temperature and vibrations during transport. Rapid changes in temperature resulted in contraction of the collagen and detachment from the tube, causing the tumoroid to float on top of the medium. This was extremely undesirable, as it would compromise the stability of the samples and the ability to image them with

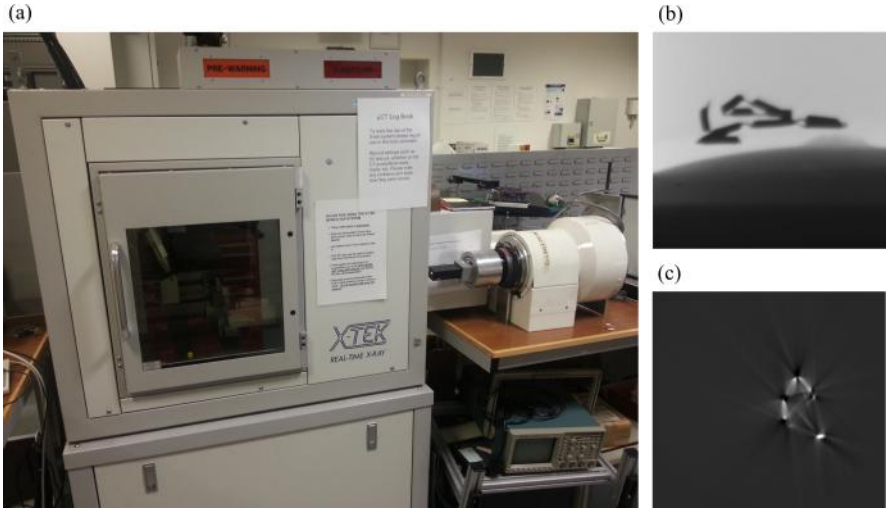

Figure 8.14: *(a) X-TEK Real-time x-ray benchtop μCT, and images acquired of Gold Anchor^{TM}: (b) projection and (c) reconstructed 3D image.*

Figure 8.15: *Polystyrene box for transport of the tumoroids between the Royal Free Hospital and UCL main campus.*

the optimised setup. To avoid this issue a prototype of a transportation box was designed and produced in order to minimise the perturbation of the samples during relocation. It consisted of polystyrene box for temperature isolation, with slots for two samples.

8.4.3.5 Timeline for imaging sessions

Figure 8.16 presents the general timeline followed for the imaging sessions. The samples were produced approximately 24 hours before the time the MR system was booked for. The whole process took approximately 2 hours per sample, followed by the incubation period. In the day following the production, the samples were transported to CABI using the adequate transportation box. The bijou tube was only fully filled with medium right before transportation, to minimise the risk of its weight causing the collagen to migrate to the top of the tube. The imaging sessions lasted, in general, 4 to 5 hours, with T_1 and T_2 maps being acquired. The first hour is occupied with setup,

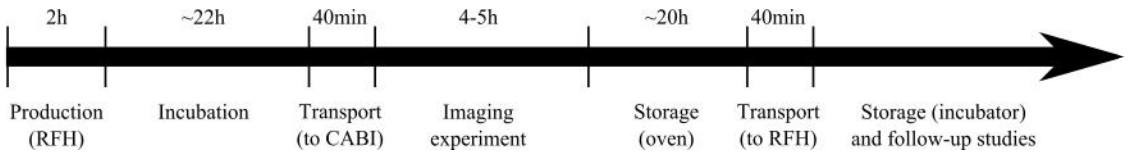

Figure 8.16: *Timeline followed in the imaging sessions, from production to imaging experiments to follow-up studies.*

tuning of the RF, scout and localiser acquisitions, and preliminary acquisitions to tune any parameters deemed necessary. During the imaging, the temperature is stabilised in a warm water bath. If more than one sample was transported, this sample is stored in the oven while the first is being imaged. Once the first imaging session is completed, the two samples switch places. After all images are acquired, the samples stay in the oven until returned to the Royal Free Hospital. This can be done on the same day if follow-up studies of the samples are planned.

8.5 Magnetic resonance imaging of the tumoroids

This section describes the several imaging experiments performed with the tumoroids. These studies were conducted in chronological order, as in this preliminary work each session raised different research questions that following experiments attempted to answer.

8.5.1 Methods and materials

8.5.1.1 Samples description

The tumoroid samples were manufactured by following the standard protocol to seed cells in collagen hydrogel (described in detail in Section 8.2) that was modified to further increase the cell density of the construct. The plastic compression process was performed during 40 seconds, and additional weight was applied after that time. The whole (or partial) dense ACM was then immersed in a uncompressed collagen matrix. Some of the samples imaged also contained different types of markers (previously discussed in section 8.4.3.3).

A total of three sequential imaging studies were performed, with different samples being imaged sequentially for different aims (Figure 8.17). All the sample were manufactured by Tong Long (Division of Surgery and Interventional Science, UCL) as per my design specifications.

Study I: Two samples with complete compressed constructs were produced: one was acellular (0M), and the second was seeded with $30{\times}10^6$ HT29 cells (30M). Figure 8.18

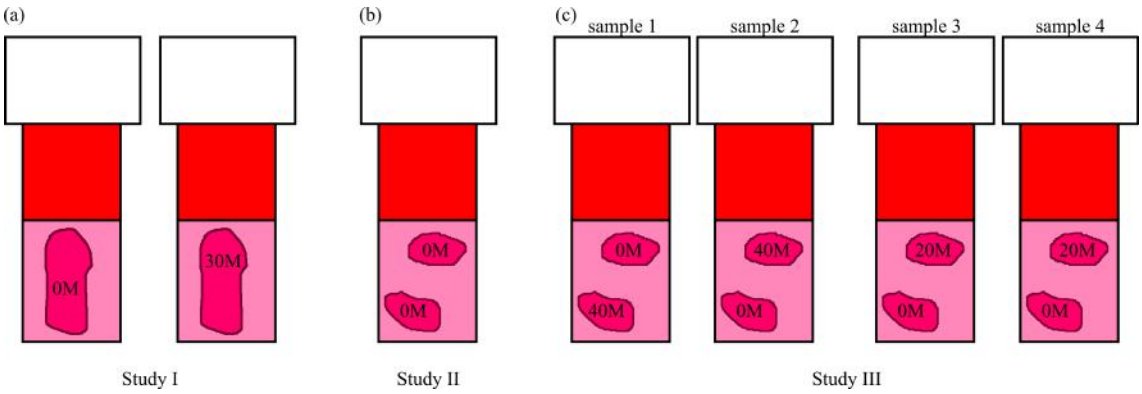

Figure 8.17: *(a) Schematic diagram of the samples used for the imaging studies: (a) Study I: acellular and 30×10^6 cells, (b) Study II: acellular, and (c) Study III: acellular and 20×10^6 or 40×10^6 cells.*

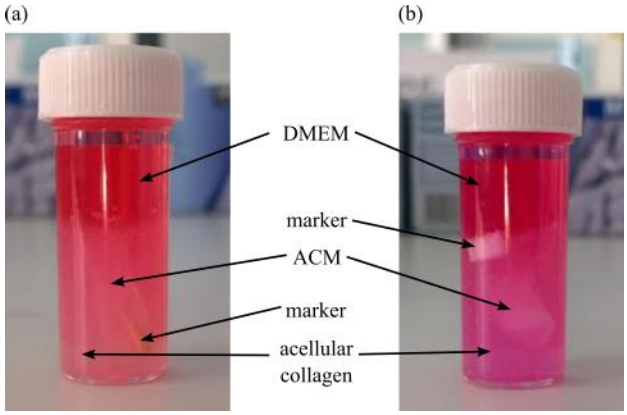

Figure 8.18: *(a) Acellular (0M) and (b) seeded with 30×10^6 HT29 cells (30M) tumoroid samples. Different markers were tested with these samples.*

shows the two samples before imaging. The aim was to verify the ability of generating MR contrast with the tumoroids.

Study II: One acellular sample (0M) composed of the two halves of a compressed construct was produced. The aim was to evaluate the MR contrast generated by compressed collagen in the absence of cancer cells.

Study III: Two types of samples with two halves of compressed constructs were produced: the first consisted of an acellular portion (0M) and a 20×10^6 HT29 cells portion (20M), and the second consisted of an acellular portion (0M) and a 40×10^6 HT29 cells portion (40M). Each type of sample was produced twice (with the 40M ones being produced from different collagen neutralisation, and the 20M coming from the same neutralisation split in half), with the relative positioning of each portion being switched. The aim was to evaluate the sensitivity of MR acquisitions to changes in biology via measuring changes in image intensity versus cell seeding density.

8.5.1.2 Data acquisition

The Bruker ICON^{TM} MRI system for pre-clinical research was used to image the samples. Standard standard T_1, T_2 and T_2^* sequences were acquired for each sample.

The samples were imaged alive, and kept in the oven at 35° when not being imaged. During the imaging sessions the temperature was monitored and kept above 32° to minimise loss of integrity throughout the imaging sessions.

The 3D printed sample holder described in Section 8.4.3.2 was used from study II onward. For study I the bijou tube was fitted inside the coil using support material, and therefore the filling was not fully optimised for this experiment.

8.5.1.3 Measurement of T_1 and T_2 relaxation times

To measure the T_1 and T_2 relaxation times, the imaging data is fit to equations 8.3 and 8.4. There are two different ways to process this data: a ROI-based approach, or a pixel-by-pixel fitting. In the first case the average signal over a ROI is used, while for the second an histogram of relaxation times is used to find the peaks in T_1 (or T_2) correspondent to the different constituents of the sample.

Measured signals differ from the true signal due to the existence of noise. Due to the Rician properties of the MR noise, M_{xy} is not measured as decaying to zero but rather to an offset value. The noise in the images is a systematic factor, rather that just a source of random variation. The relationship between measured and true signal can be well approximated using a simple function:

$$S = \sqrt{S_0^2 + C^2} \tag{8.5}$$

where S is the measured signal, S_0 the true signal in the absence of noise and C the noise-related constant [255]. C can be estimated by using the mean signal intensity in a region devoid of true signal.

Therefore, in T_2 measurement experiments equation 8.4 is modified to include the noise properties:

$$S = \sqrt{(S_0 e^{-TE/T_2})^2 + C^2} \tag{8.6}$$

For T_1 measurements another important concept is the efficiency of the IR process. When the inversion is perfect, the signal at TI=0 should be equal to -M_0 [256]. Therefore, equation 8.3 was modified to include loss of efficiency during the IR process and the effect of noise.

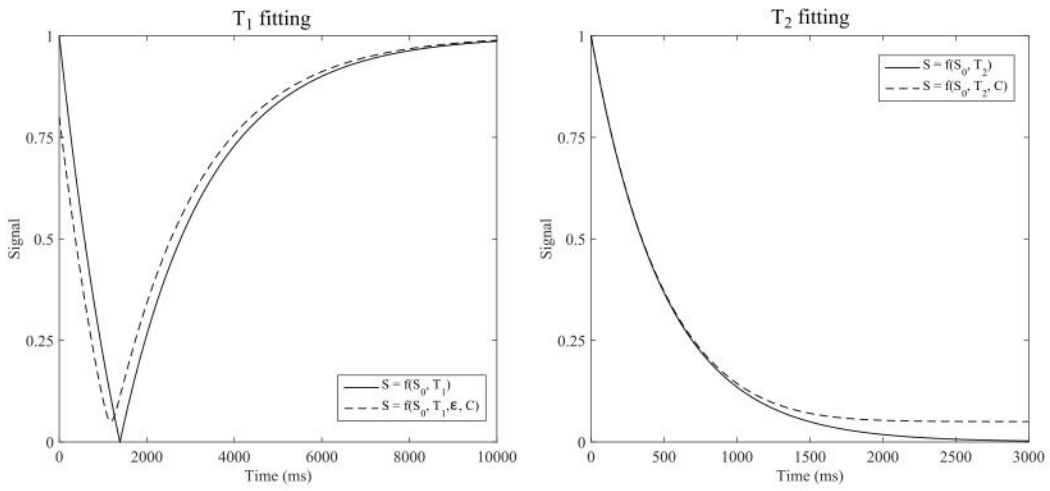

Figure 8.19: *Comparison between T_1 and T_2 fitting curves: theoretical vs experimental (considering the noise).*

$$S = \sqrt{[S_0(1 - \epsilon \times e^{-TI/T_1})]^2 + C^2} \tag{8.7}$$

where ϵ represents the efficiency of the inversion recovery process.

The differences between the original equations and the ones used during the fitting of experimental data can be schematically evaluated in Figure 8.19.

The code to estimate the T_1 and T_2 relaxation times was implemented in MATLAB (MathWorks, Natick, MA, USA), using the curve fitting tool (non-linear least squares method and the trust-region algorithm).

8.5.2 Results and discussion

8.5.2.1 Study I

The aim of this experiment was to evaluate the feasibility of generating contrast with the tumoroids in standard MR acquisitions, and to generate feedback for next iterations of the project in terms of experimental setup. The imaging session was split in two halves: first the acellular tumoroid was used to optimise the setup and to perform some initial tuning of the imaging parameters. Later it was replaced with the 30M tumoroid to repeat and improve the acquisitions.

The acellular tumoroid sample was imaged with scout T_1, T_2 and T_2^* sequences. The scout T_1 and T_2 provided evidence of structural information, and were extended to sequences that measure T_1 and T_2 relaxometric maps.

T_1 contrast was investigated using an IR-RARE sequence with TE=12 ms, TR=10000 ms, and TI={500, 1500, 2000, 2500} ms, with a 20×20 mm^2 FoV and resolution of

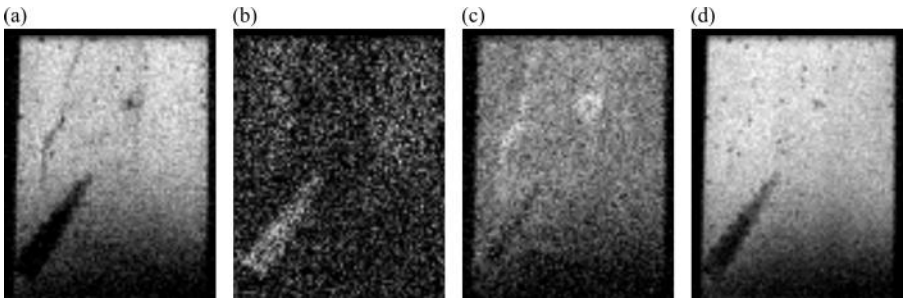

Figure 8.20: *IR-RARE sequences images of an acellular tumoroid with TI of (a) 500, (b) 1500, (c) 2000, and (d) 2500 ms (TE=12 ms, TR=10000 ms).*

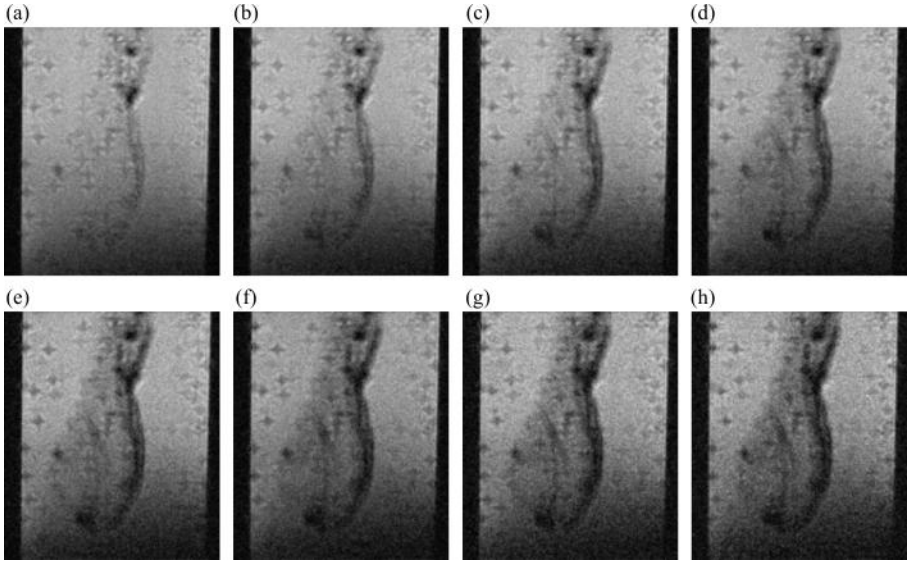

Figure 8.21: *T_2 MSME images of an acellular tumoroid with a TE of (a) 50, (b) 100, (c) 150, (d) 200, (e) 250, (f) 300, (g) 350 and (h) 400 ms (TR=6000 ms).*

0.208×0.208×1 mm^3 (Figure 8.20). The images show contrast at the boundaries, but no apparent signal difference between the dense collagen (internal) and the uncompressed collagen (external).

T_2 maps were assessed using multi-slice multi-echo (MSME) technique with TE=25:25:400 ms (a total of 16 echoes) and TR=6000 ms, with a 20×20×14 mm^3 FoV and resolution of 0.125×0.125×1 mm^3 (Figure 8.21). The images show evidence of contrast particularly at the boundaries, and (but less evident) between the dense collagen (internal) and the uncompressed collagen (external). The air bubbles formed in the uncompressed collagen generated artefacts particularly in T_2 images.

T_2^* scout images were acquired with two FLASH sequences, with the following parameters: TE=5 ms, TR=400 ms and α=30°, and TE=40 ms, TR=800 ms and α=30°, with a 20×20×14 mm^3 FoV and resolution of 0.156×0.156×1 mm^3 (Figure 8.22). It is clear that the air bubbles present in the collagen gel generate severe artefacts in T_2^* images, and therefore the current tumoroid model was not adequate for T_2^* imaging.

The 30M tumoroid sample was also imaged to measure T_1 and T_2 relaxation times.

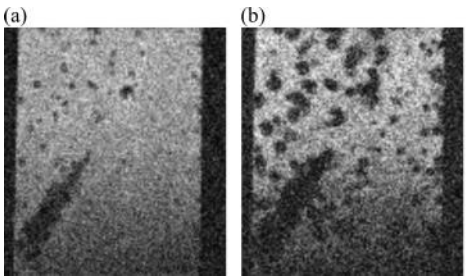

Figure 8.22: *T_2^* FLASH sequences images of an acellular tumoroid with (a) TE=5 ms, TR=400 ms, α=30° and (b) TE=40 ms, TR=800 ms, α=30°.*

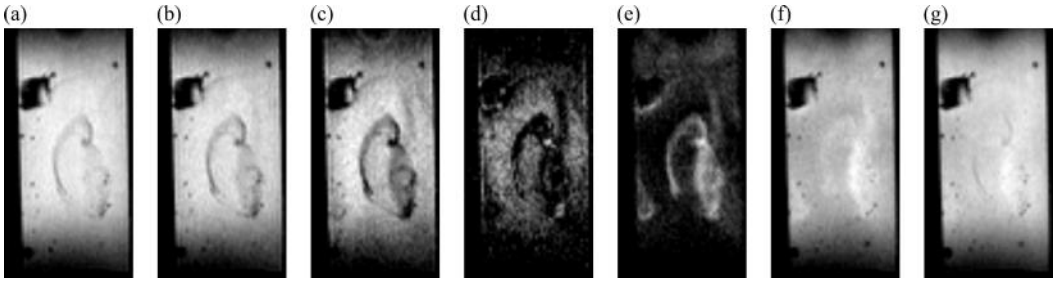

Figure 8.23: *IR-RARE sequences images of a 30M tumoroid with TI of (a) 205, (b) 500, (c) 1000, (d) 1500, (e) 2000, (f) 4000, and (g) 6000 ms (TE=12 ms, TR=10000 ms).*

T_1 maps were assessed using an IR-RARE sequence with TE=12 ms, TR=10000 ms, and TI={205, 500, 1000, 1500, 2000, 4000, 6000} ms, with a 32×32 mm^2 FoV and resolution of 0.333×0.333×1 mm^3 (Figure 8.23). T_2 maps were assessed using the MSME technique with the same parameters as the previous sample but a 32×32×14 mm^3 FoV and resolution of 0.333×0.333×1 mm^3 (Figure 8.24).

In both sequences appears to exist a difference in signal between the dense collagen and uncompressed collagen, but a striking feature is the "halo" effect at the boundaries of the ACM. The signal intensity profiles clearly indicate a dip in signal intensity at the boundaries (Figure 8.25).

This first study showed the ability to generate contrast in the ACM using MRI. However, the source of the contrast and the origin of the "halo" were uncertain. There are three possible sources of contrast: cells, compressed collagen, and air bubbles. Air bubbles at the boundaries could cause apparent contrast, but this possibility was discarded as (i) the volume of the region was not affected by increasing the TE of the acquisitions and (ii) that region signal was effectively inverted during the IR. Therefore, either gradients of collagen or cell density (or combination of both) cause the difference in signal. The "halo" could be also explained by either, as the collagen is more compressed at the boundaries and cells tend to migrate from the hypoxic center with time. However, it is more likely to be collagen generated, as visual inspection of the tumoroids immediately after manufacturing shows regions at the boundaries of stronger white colouring. Contact between collagen and metal, and stronger forces felt at the boundaries during plastic compression can explain the additional compression of the collagen.

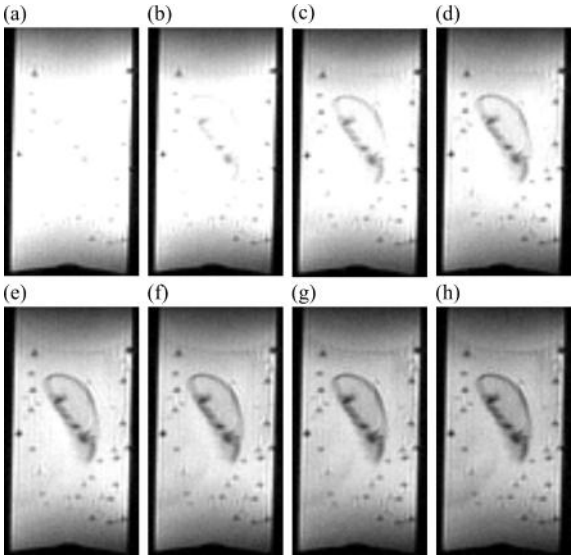

Figure 8.24: *T_2 MSME images of a 30M tumoroid with a TE of (a) 50, (b) 100, (c) 150, (d) 200, (e) 250, (f) 300, (g) 350 and (h) 400 ms (TR=6000 ms).*

The acquisitions in this study were purely qualitative and were not adequate for further quantitative analysis. However, this preliminary study was fundamental to optimise the setup and imaging acquisitions. The 3D printed sample holder described in section 8.4.3.2 was produced based on the feedback obtained from Study I.

8.5.2.2 Study II

As a result from the previous study, the sample holder described in section 8.4.3.2 was designed and 3D printed. It was therefore necessary to assess the suitability of the holder, and if its current dimensions were adequate.

The sample used to test the experimental setup was acellular, and the whole construct was split and separated in two parts. The portions were immersed in the less collagen, such that one was on top of the other, and as close to the bottom/top of the surrounding collagen. This disposition allowed to clearly distinguish the two ACMs (without the need of fiducial markers) and to evaluate the positioning of the ACMs inside the imaging FoV to ensure that the regions to image were well positioned inside the coil (i.e., within volumes of larger signal received). The sample was acellular to facilitate its production and provide further preliminary results into the mechanism of contrast generation (i.e., infer if compressed collagen had different MR properties than uncompressed collagen).

The acellular tumoroid sample was imaged to measure T_1 and T_2 relaxation times. T_1 maps were assessed using an IR-RARE sequence with TE=12 ms, TR=10000 ms, and TI={250, 500, 750, 1000, 1250, 1500, 1750, 2000, 4000, 6000, 8000} ms, with a 40×40 mm^2 FoV and resolution of 0.42×0.42×1 mm^3 (Figure 8.26). T_2 maps were assessed using MSME technique with TE=25:25:3200 ms (a total of 128 echoes), with a 40×40 mm^2 FoV and resolution of 0.25×0.25×1 mm^3 (Figure 8.27).

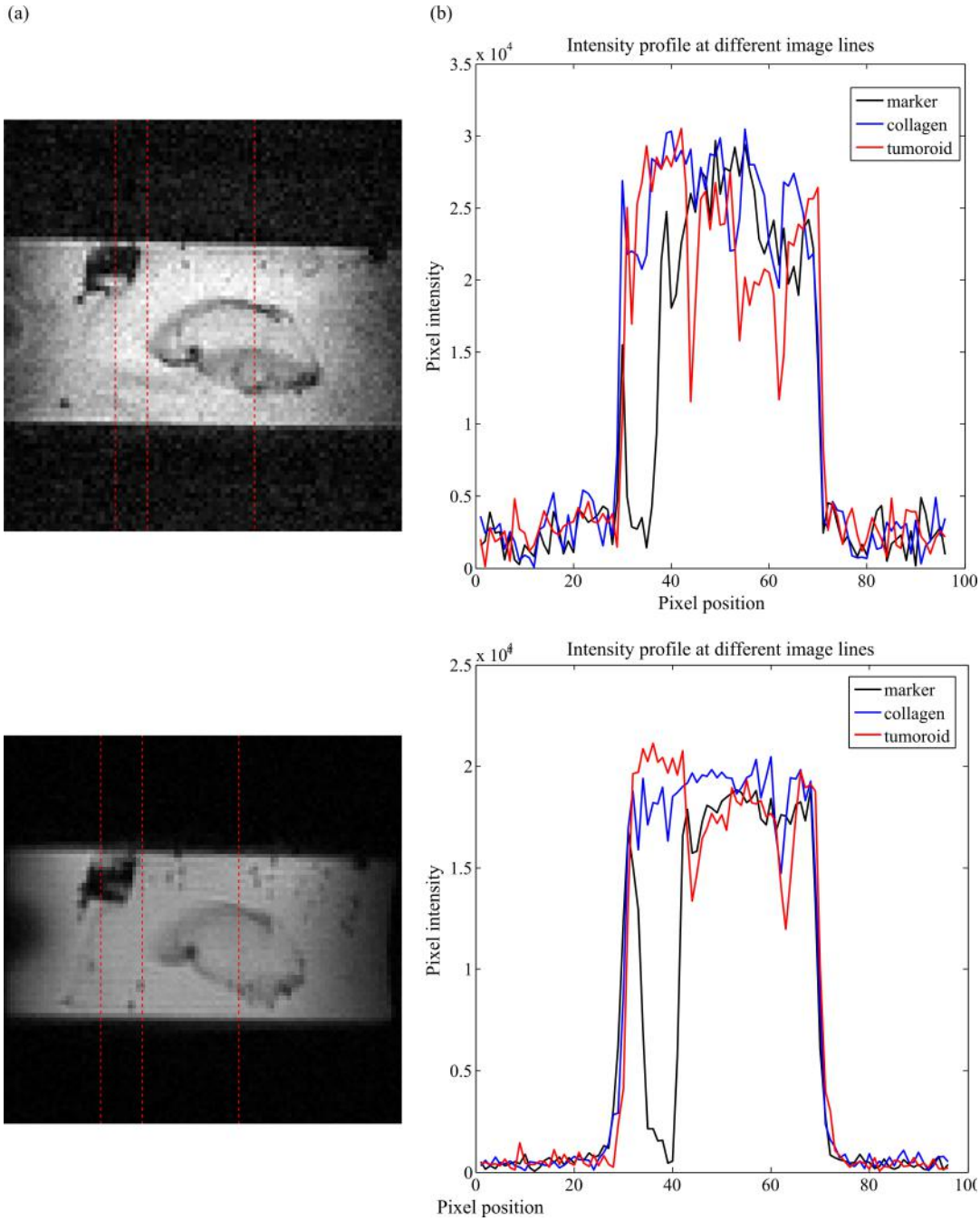

Figure 8.25: *Intensity profile on T_1(TI=500 ms) (top row) and on T_2 (TE=300 ms) (bottom row) images: (a) lines in the image where the profile was plotted and (b) intensity profile plot.*

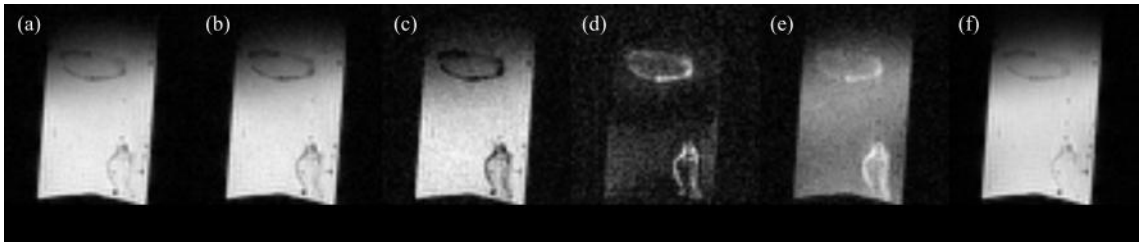

Figure 8.26: *IR-RARE sequences images of an acellular tumoroid (2×) with TI of (a) 250, (b) 500, (c) 1000, and (d) 1500 ms, (e) 2000 ms, and (f) 8000 ms (TE=12 ms, TR=10000 ms).*

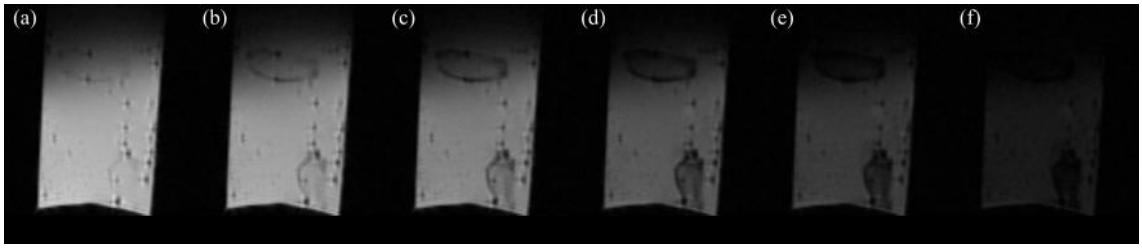

Figure 8.27: *T_2 MSME images of acellular tumoroid (2×) with a TE of (a) 100, (b) 200, (c) 400, (d) 800, (e) 1200, and (f) 2000 (TR=5000 ms).*

Unlike in the previous study, the overall morphology of the acellular tumoroid is similar to that found for the cellular sample used in study I (i.e., reduced intensity within the ACM, and a darker "halo" at the construct boundaries). The differences between the results of study I and II regarding the acellular gel indicate the poor reproducibility of the samples. It is also clear that compressed collagen is one of the sources of the contrast, but further studies are necessary to evaluate if the cells also contribute to the final intensity. Therefore, the samples produced for following experiments were produced in a more reproducible fashion, using the protocol described in Figure 8.9.

Regarding the sample holder prototype, while both ACMs are clearly captured within the imaging FoV (Figures 8.26 and 8.27) the upper ACM was partially within the volume where the signal intensity starts to decay, while the region inferior to the bijou tube still had stronger signal. Therefore design of the holder was afterwards adjusted to reflect these findings, and a new version was printed for study III.

8.5.2.3 Study III

The aim of this study was to evaluate the role of the HT29 cells into the contrast generation of the tumoroids. To achieve this goal ACMs of different cell seeding density (0, 20 and 40 $\times 10^6$) were imaged.

All the samples were imaged to measure T_1 and T_2 relaxation times. T_1 maps were assessed using an IR-RARE sequence with TE=12 ms, TR=10000 ms, and TI={250, 500, 750, 1000, 1250, 1500, 1750, 2000, 4000, 6000, 8000} ms, with a 40×40 mm^2 FoV and resolution of 0.42×0.42×1 mm^3 (Figure 8.28). T_2 maps were assessed using the MSME technique with TE=25:25:3200 ms (a total of 128 echoes), with a 40×40 mm^2 FoV and resolution of 0.25×0.25×1 mm^3 (Figure 8.29).

Visual inspection of the imaging data acquired shows that the different samples all had very similar properties, and the existing differences were not perceptible with visual inspection. The most interesting aspect to note is the variability between and within samples. For example, the acellular tumoroid portion of samples 1 and 2 had very different properties. Similarly, the samples can have internal features easily distinguishable. This indicates (a) low reproducibility of the samples production and (b) the existence of

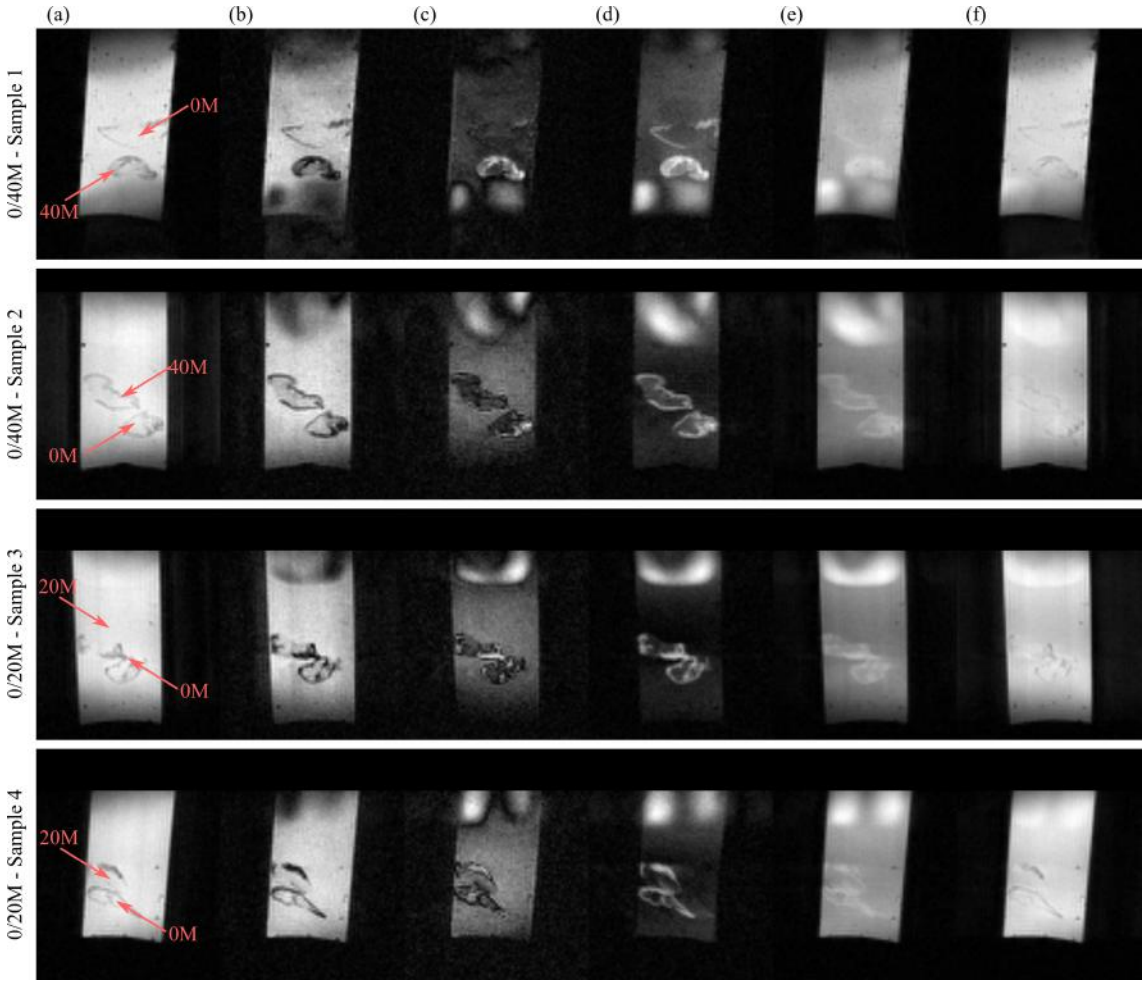

Figure 8.28: *IR-RARE sequences images of the different samples imaged with TI of (a) 250, (b) 1000, (c) 1500, (d) 2000 ms, (e) 4000 ms, and (f) 8000 ms (TE=12 ms, TR=10000 ms).*

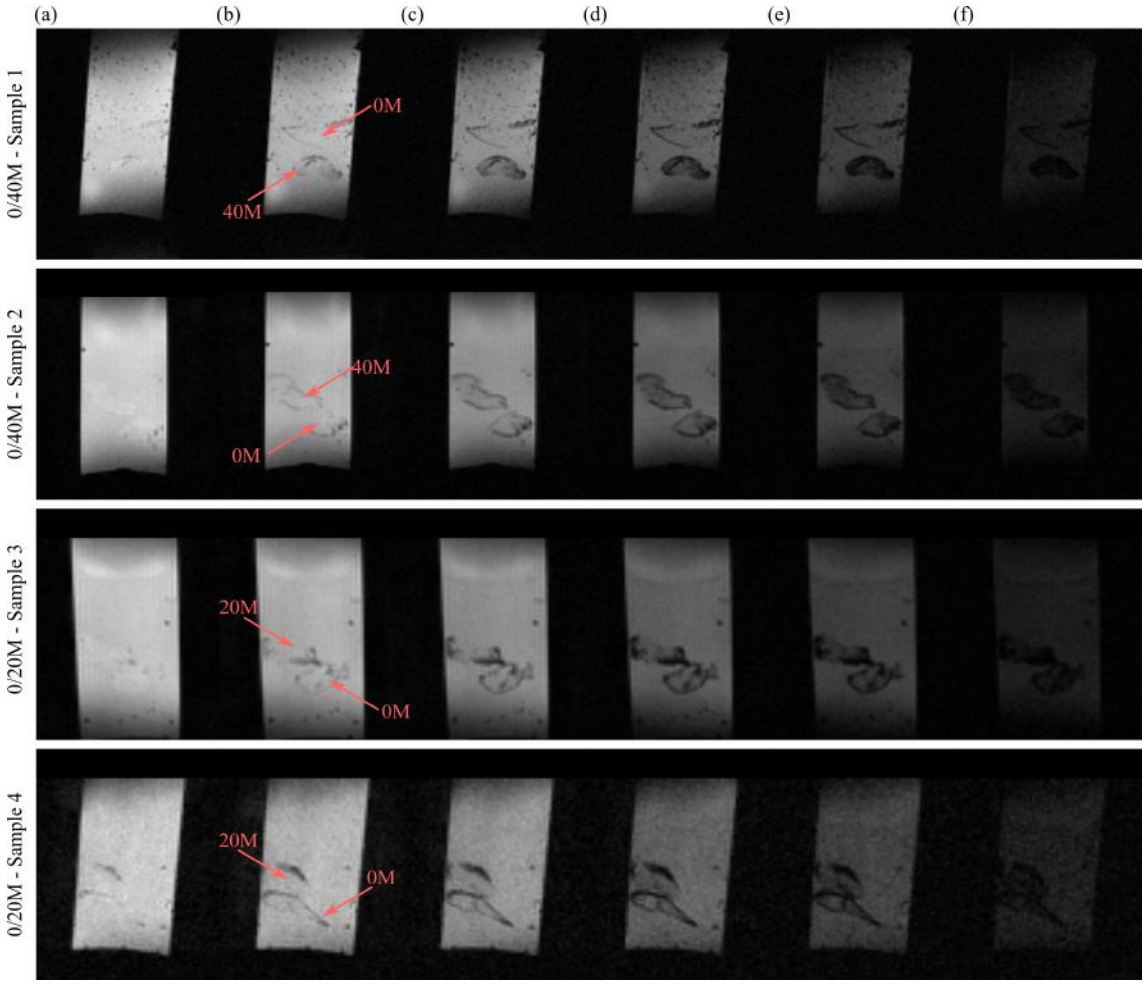

Figure 8.29: *T_2 MSME images of the different samples imaged with a TE of (a) 100, (b) 200, (c) 400, (d) 800, (e) 1200, and (f) 2000 ms (TR=5000 ms).*

gradients of compression within the samples (i.e., non-uniformity).

Using portions of the whole construct may reduce the uniformity of the samples, as additional handling (such as pinching) may deform and change locally the compression of the samples. Smaller ACMs were investigated to facilitate parallel biological studies of the samples, and were in fact adequate regarding the resolution of the MRI scanner. However, for future characterisation studies it is advisable to use the whole construct instead.

8.5.2.4 Measurement of T_1 and T_2 relaxation times

In the previous sections an overview of the experiments and images acquired was presented in a qualitative fashion. Here a quantitative analysis of the imaging data is performed, aiming to measure the T_1 and T_2 relaxation times of the tumour models.

To compute the T_1 and T_2 the noise of the data was first pre-processed. For T_1 maps, since each TI corresponds to a different acquisition, the signal should be normalised to the noise level before the fitting. For T_2 this is not always necessary as the whole range of TE is acquired in a train of SE sequences and, therefore, the noise conditions should be similar. For each dataset the properties of the noise were assessed, and corrected for when deemed necessary. The noise level was defined as the average noise within a ROI devoid of true signal. This region had to be properly defined per dataset and acquisition, as it is usual for non-signal areas to be polluted with signal from other regions.

The relaxations times were calculated using both the ROI and pixel-by-pixel approaches. Regions of interest were drawn on the images acquired for the different studies defining the surrounding uncompressed and acellular collagen, and the dense collagen (populated or not with cells). The samples imaged in study I were not included in this analysis as the acquisitions were sub-optimal for this purpose. The results obtained for T_1 and T_2 relaxation times using the ROI approach can be found in Table 8.2. The quality of the fittings can be seen in Figure 8.30. Regarding the T_1 results, there was a large variability within similarly designed samples; for example, considering the stroma the values range within an interval of [2450, 3200] ms. However, the corresponding results for T_2 were stable and reproducible. Considering different types of samples, uncompressed collagen was clearly different from compressed collagen, but the differences found between varying cell seeding densities were not significant due to the high variability within similar samples.

The ROI fitting provides an overall view on T_1 and T_2 inter-sample variability, but is an oversimplification due to the existence of clear features within the samples, such as the "halo" effect (intra-sample variability). Figure 8.31 shows the results for the pixel-by-pixel approach, which provides a qualitative view into the homogeneity of the samples. It is therefore possible to use an histogram to describe the distribution of T_1 and T_2 values

Table 8.2: *T_1 and T_2 relaxation times for stroma (acellular and uncompressed collagen) and compressed collagen seeded with 0, 20 and 40 $\times 10^6$ HT29 cells.*

	T_1				T_2			
	Stroma	0M	20M	40M	Stroma	0M	20M	40M
Study II	2430	2140	-	-	970	580	-	-
Study III								
(sample 1)	2600	2460	-	1860	970	810	-	500
(sample 2)	2980	2350	-	2670	970	680	-	680
(sample 3)	3190	2530	2560	-	1020	710	640	-
(sample 4)	3170	2640	2690	-	1000	660	670	-
Mean	2900±300	2400±200	2620±90	2300±600	990±20	690±80	660±20	590±130

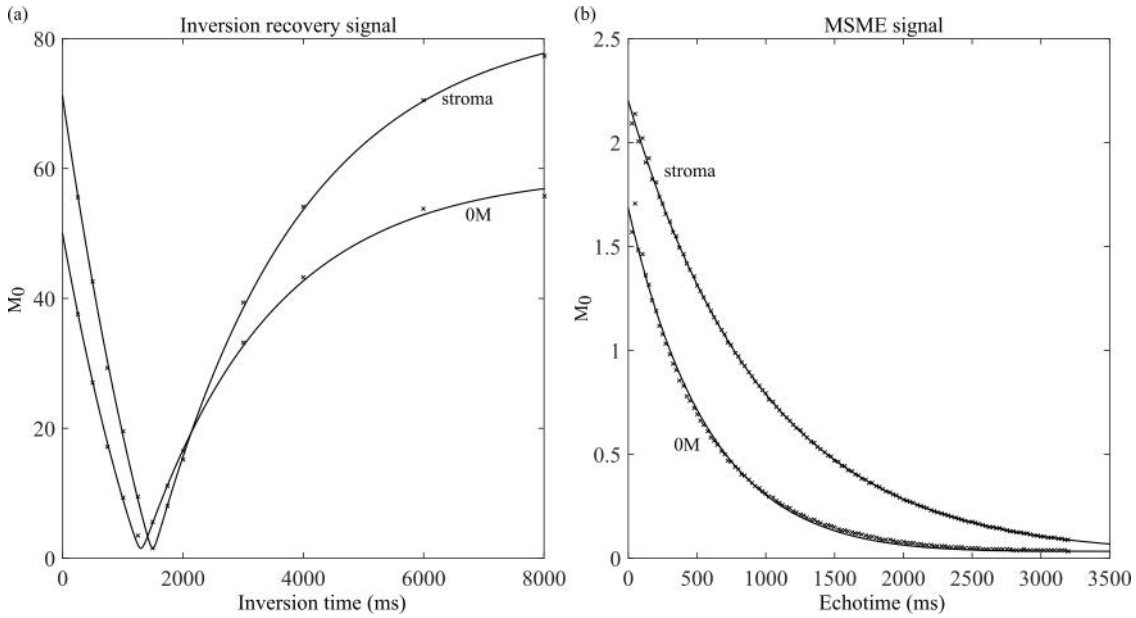

Figure 8.30: *Fitting of (a) T_1 from inversion-recovery sequence and (b) T_2 from MSME images for stroma and acellular compressed collagen (study II). The fitting was done over the average intensity within a region-of-interest.*

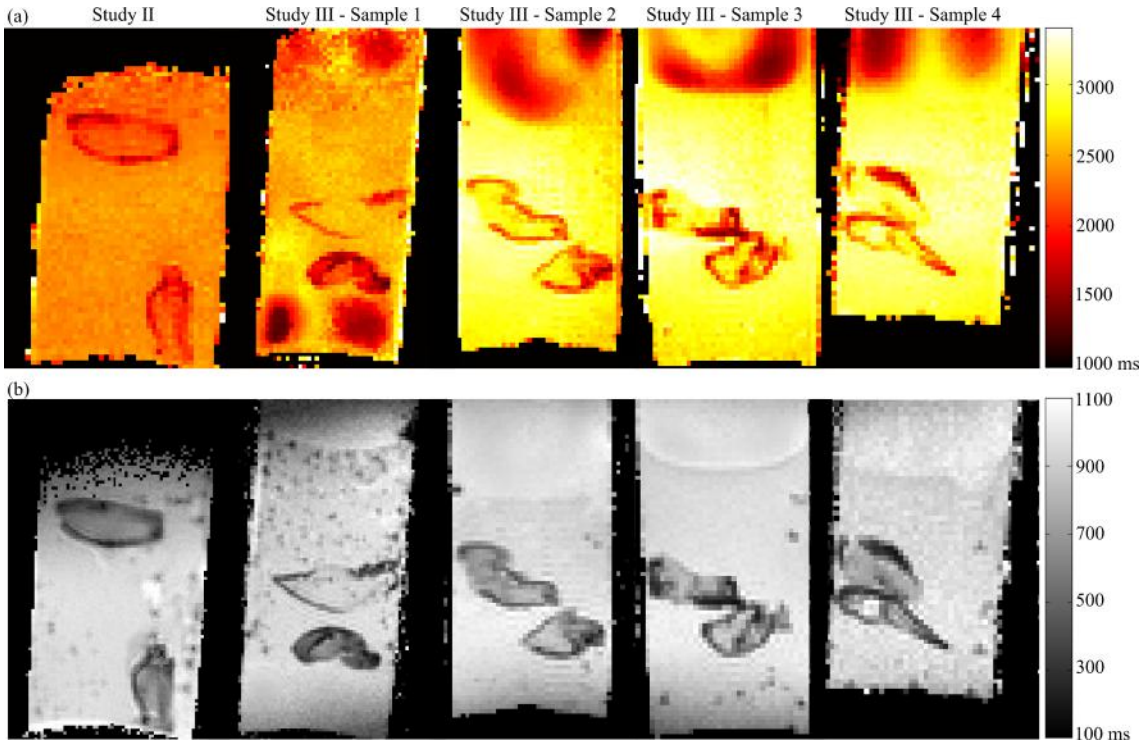

Figure 8.31: *(a) T_1 and (b) T_2 relaxometric maps of the samples imaged in studies II and III.*

(Figures 8.32 and 8.33). These figures illustrate the lack of homogeneity and reproducibility within compressed collagen samples, with values of T_1 and T_2 ranging over a wide interval of values. The variability between similar samples imaged at different time points may be related with the scanner acquisitions, such as temperature fluctuation and/or drift of the magnetic field. However, the most important effect is the poor reproducibility of the plastic compression process. A pos-imaging follow-up study of the samples used in study III showed that the collagen density varied between 10 to 80%, adding evidence of the poor reproducibility between samples.

Quantifying T_1 and T_2 allows to not only quantitatively differentiate different tissues, but also optimise TR and TE for T_1 and T_2 image acquisition protocols. The highest contrast achievable between two materials with different T_1 is achieved by selecting a TR such that the absolute difference between the signals (using equation 8.1) is maximum. The same is true for T_2 and choice of TE based on equation 8.2. Figure 8.34 shows this calculation for the average values found for the uncompressed collagen and 40M sample. In this case the optimal values for imaging were TR=2580 ms and TE=760 ms.

8.6 Current status and future work

Considering the design specifications of an imaging bio-phantom (identified in section 8.4.1), at the current stage of this project the following progress was achieved:

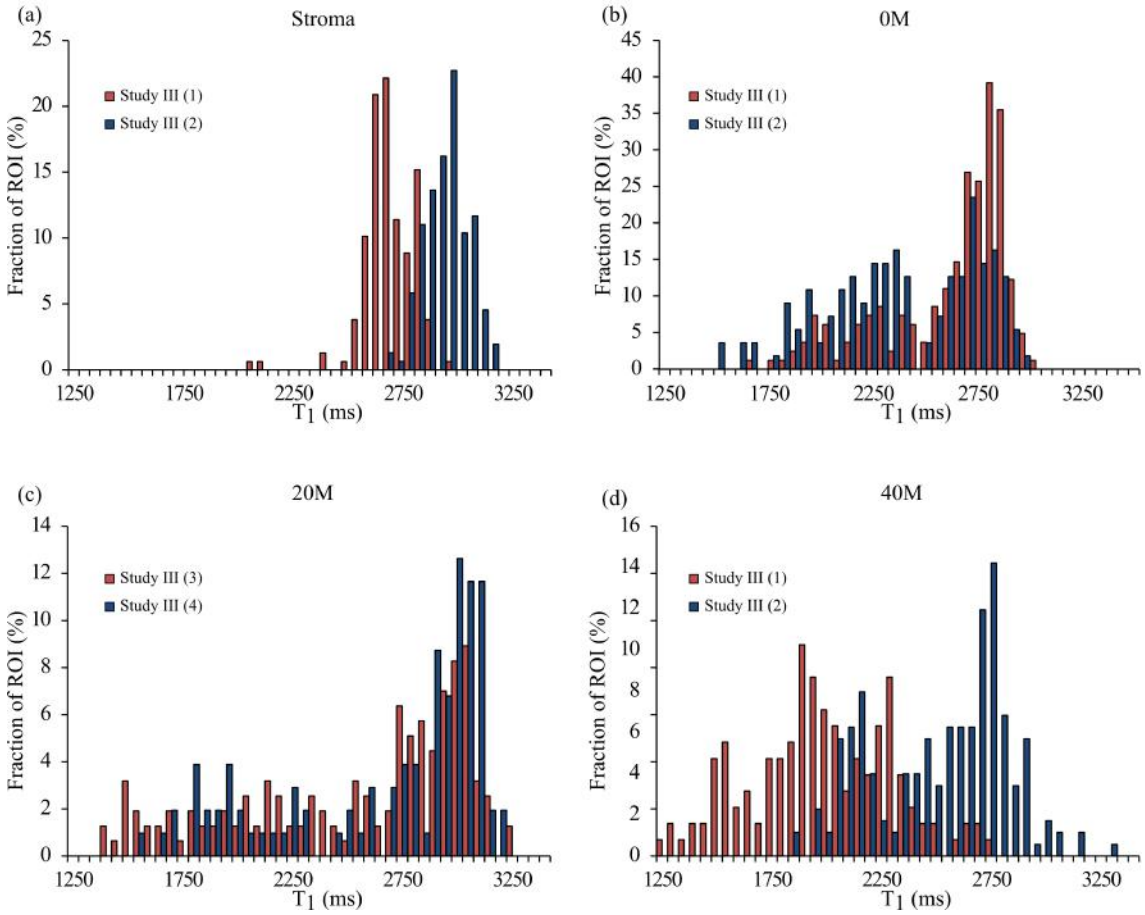

Figure 8.32: *T_1 histograms for (a) stroma (uncompressed collagen), and compressed collagen seeded with (b) 0, (c) 20×10^6 and (d) 40×10^6 HT29 cells.*

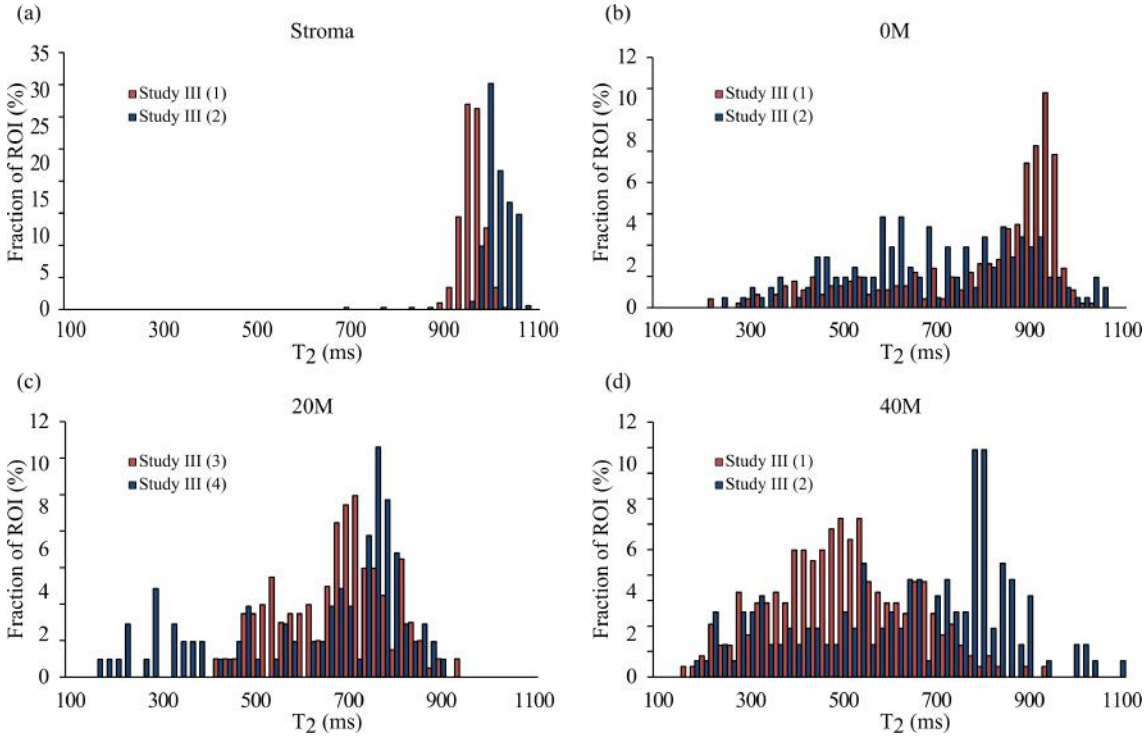

Figure 8.33: *T_2 histograms for (a) stroma (uncompressed collagen), and compressed collagen seeded with (b) 0, (c) 20×10^6 and (d) 40×10^6 HT29 cells.*

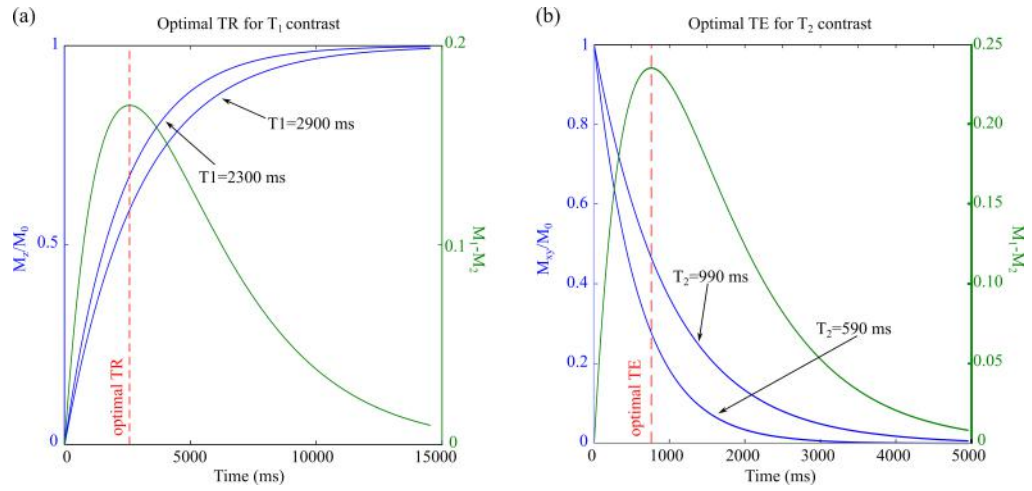

Figure 8.34: *Optimal acquisition values for (a) T_1 and (b) T_2 contrast for uncompressed collagen and compressed collagen seeded with 40 $\times 10^6$ cells.*

(i) *The tumoroids should provide image contrast in different imaging modalities.* The tumoroids were engineered to be imageable using T_1 and T_2 sequences. The current model was found inadequate in its current state for T_2^* and CT imaging. Other multiparametric acquisitons (such as DW) were not investigated at this point.

(ii) *The contrast generated should be quantifiable based on its biological properties.* The non-homogeneity and poor reproducibility of the samples did not allow to take firm conclusions on the contrast mechanisms, and the current model does not generate controllable and quantifiable contrast.

(iii) *The tumoroids should be compatible with sequential imaging.* The samples were imaged alive, and the viability of the cells was confirmed via confocal microscopy several days after the imaging sessions, providing evidence of the suitability of the model for sequential imaging studies. Further studies are necessary to ensure resistance to multiple travelling and characterise the biological strain due to transport and imaging.

(iv) *Changes in the biological properties of the tumoroids should be within the sensitivity of the imaging system.* A quantifiable phantom has to be achievable before this specification can be evaluated, and therefore this aspect is unexplored to this point.

To achieve the design requirements, three lines future of future research were identified:

1. Sample production: design, reproducibility and engineering.

2. Characterisation and biological properties of the samples.

3. MRI setup.

8.6.1 Sample production: design, reproducibility and engineering.

The artificial cancer mass bio-phantom needs to be further engineered to be increasingly specialised for imaging applications. The current protocol to produce the samples is considered reproducible from a tissue engineering point-of-view, but not from a physics point-of-view. Design needs, possible improvements and other aspects of potential interest for future research that were not further explored in this stage of the project are discussed in the following paragraphs:

- The non-homogeneity and poor reproducibility is the most important feature to understand, engineer and characterise. There are several aspects to consider in this point:

  (i) The setting of a collagen gel is dependent on the pH of the solution, and there is a range where the process is stable. This leads to variability in the samples produced, as the chemical properties can vary within that interval. The protocol for neutralisation is not exact, as the base is added on a drop-wise fashion. Therefore, a step that can be further optimised is the pH neutralisation. Since there can be variability between batches of collagen, a method to measure quantitatively the pH in real-time should be investigated and developed.

  (ii) The method currently used for plastic compression is clearly not reproducible and controllable between batches of samples, as seen in the imaging and collagen density studies. The plastic compression introduces gradients of collagen density that are complex to model theoretically. Other methods to further compress the gel could be investigated further, such as compressing the tumoroids to thin sheets, and rolling the structures. Alternatively, the use of uncompressed collagen can be explored to remove the uncertainty in matrix density. However, this also has limitations, such as less similarities with *in vivo* tumours, larger number of cells being necessary to increase cell density, and imaging artefacts caused by air bubbles.

  (iii) The possibility of generating heterogeneities and internal features in the bio-phantoms is unwanted when attempting to quantify the biological properties behind the contrast generated, but of interest for image registration validation studies so the possibility of designing samples with internal features is of interest. Therefore, two routes can be followed: design a homogeneous tumoroid, with controllable cell and matrix densities, or develop other imaging methods method to quantitatively measure these two quantities and coregister that information with MR imaging, so that it is possible to correlate signal intensity with the gold-standard information.

- The process of setting the collagen gel results in the production of air bubbles inside the matrix, that generate artefacts in MR images. A method to minimise bubble

production needs be engineered to minimise imaging artefacts.

- Different cell lines have different micro-structure (i.e., size, shape and arrangement in the matrix) and biological responses. Different cell lines can be used to investigate how the properties of the underlying microstructure affect the measured MR signal.

- Previous studies showed that when the tumoroid is co-cultured with other types of cells (e.g., fibroblasts and endothelial cells) the tumour cells showed different morphology and behaviour than when cultured alone in the collagen gel. When the ACM is surrounded by collagen gel populated with other cell lines the cancer cells have slower proliferation than when cultured alone [235]. Acellular gels were used in the described imaging experiences; nevertheless, imaging of more complex stroma models is of potential future interest.

- Contrast agents may also be included in the current model to enhance contrast [244, 257]. Ideally the model should be usable without external contrast agents, but it might be that engineering of the constructs is not enough to achieve the design specifications and artificial contrast is necessary. Nevertheless, there is a lot of potential in studying the suitability of different contrast agents using this tumour model for other applications, such as the development of biomarkers.

8.6.2 Characterisation and biological properties of the samples

Together with improving the reproducibility and controllability of the tumour model, it is fundamental to understand how the changes in design and the imaging experiments affect the biological properties of samples. Several points can be highlighted in this topic:

- The tumoroids are living samples, and their micro-structure changes in time due to cell proliferation and migration. For sequential imaging studies, it is necessary to understand over time how the cells are distributed in the matrix (morphology), cell density viability (proliferation), oxygenation cell number over time (hypoxia), etc.

- Imaging experiments are a strain to the cells, as imaging experiments involve spending several hours in sub-optimal conditions (temperature, atmosphere, vibration due to transport, etc). The level of that strain must be evaluated and further minimised if deemed necessary.

- In the imaging experiments conducted a resolution of a fraction of the mm was achieved, but this is still 1-2 orders of magnitude above the dimensions of the cell/matrix (the average diameter of HT29 cells is 11 μm). It is therefore crucial to investigate other types of imaging common in biology (such as microscopy methods) and methods to register with medical imaging modalities (CT/MR).

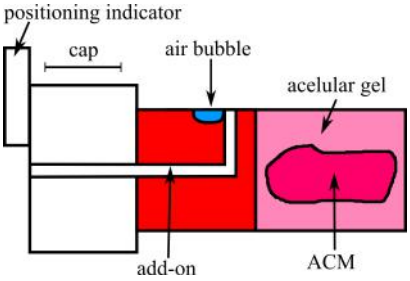

Figure 8.35: *3D printer cap add-on: the add-on is glued to the cap, and when horizontal stops bubbles from reaching the collagen matrix.*

8.6.3 Magnetic resonance imaging setup

- The experimental setup and sample holder could be further refined:

  (i) An additional safety check for the setup protocol was designed its feasibility was investigated. It was not built as the previous design was proven in most cases to deal sufficiently with the air bubble issues, but it could also help dealing with potential migration of the samples within the bijou tube. This consisted of modifying the bijou cap to stop the residual air bubble from coming in contact with the collagen (Figure 8.35). This should be made of a polymer material to be easily manageable, MR-compatible and autoclavable (since it needs to be sterile). Polymers have a specific temperature (inferior to the melting point) at which the modulus drops catastrophically, and they loose their physical properties. The autoclave reaches high pressure and temperatures of 121°C, which can be problematic for several polymers used in lab supplies (Figure 8.36). The bijou container screw cap is made of polypropylene, and has a melting point of approximately 150°C. Therefore the add-on could also be made of polypropylene. Other option that facilitates manufacture is PLA, the material used by the 3D printer, which as a melting point also of approximately 150°C. The softening points are higher than the autoclave temperature, but both materials have a heath deflection temperature considerably lower (approximately 60°C). Therefore, both materials were autoclaved for testing purposes, and no major issues were seen with deformation of the constructs. Different parts can be glued together using adhesives resistant to water and high-temperatures (i.e., autoclavable). Several glues were tested and both the Permatex High-temp RED RTV silicone and the Evo-Stik Serious Glue were found to be appropriate for this application.

  (ii) The sample holder could be further engineered to maintain the samples in an ideal atmosphere during imaging (Figure 8.37), minimising the strain to the cells. The bijou screw cap could be modified to include a filter, allowing a constant airflow and minimise chance of contamination. The available screw cap can be drilled to attach a glued filter, and then autoclaved. This filter would

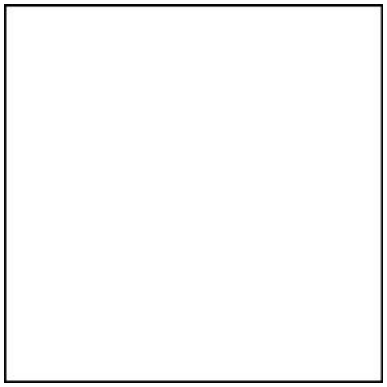

Figure 8.36: *Heath deflection temperatures under a load of 1.82 MPa for selected polymers. Adapted from [258].*

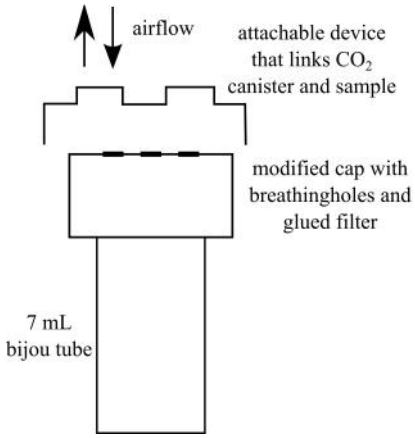

Figure 8.37: *Possible improvements to the sample holder for imaging sessions: the bijou screw cap is individualised, by drilling holes and adding an appropriate filter to keep its contents sterile. An attachable device is added to the cap to connect with a 5% CO_2, a 2% O_2 and a 93% Ar gas cannister. The aseptic and atmosphere conditions can be therefore maintained during imaging sessions.*

have to be waterproof since the tube is filled with medium. An airflow system can then be added to the cap to connect it with a 5% CO_2, 2% O_2 and 93% Ar gas cannister.

- The 1 T MRI system used in this work has its limitations. Thus, using a higher resolution system (such as the 9.4 T system also available at CABI) might improve the imaging data extracted. With higher magnetic field both larger sensitivity and resolution are achievable.

- To estimate the T_1 and T_2 relaxation times the method used can be sub-optimal in some cases, particularly when the images contain considerable noise. The non-linear least squares method assumes Gaussian noise, which is not true for MRI. To consider the Rician properties of the noise distribution the most appropriate method is the maximum likelihood fitting. Additionally, other fitting equations are also suggested in the literature that may be more appropriate [255, 256, 259].

- Other types of imaging modalities that provide complimentary structural/functional

information should be investigated, such as DW-MRI, magnetic spectroscopy, etc.

8.7 Conclusions

In this chapter the use of artificial cancer masses as a bio-phantom for multimodal and multitemporal imaging was investigated. The ideal design specifications were presented, and the investigative work conducted toward the specified goals was presented. An ideal bio-phantom should be reproducible and controllable, generate quantifiable contrast in multimodal and multiparametric imaging, and be compatible with sequential imaging. The preliminary work toward such a bio-phantom, such as design of the samples, design of the experimental setup and imaging studies were detailed in this chapter. The tumoroids were shown to be imageable in the current state using standard T_1 and T_2 MRI acquisitions. CT and T_2^* contrast was not achieved, while more complex MRI sequences were not investigated. The main limitation of the current tumour model was the poor reproducibility and controllability of the properties of the samples, which makes the contrast generated non-quantifiable.

Chapter 9

Final remarks

> The whole is more than the sum of its parts.
>
> Aristotle

This thesis has made a contribution to the Medical Physics field as a research effort focused at using information from different imaging modalities for adaptive radiotherapy applications. This was a multidisciplinary effort, where radiotherapy physics, imaging, computing and tissue engineering techniques were used toward the challenging goal of integrating multiple imaging modalities into the radiotherapy workflow toward improving patient outcome.

The majority of the work conducted in this project was focused on using DIR as a tool, and CBCT as the main image-guidance modality. Both IMRT and proton therapy applications were investigated, and the work conducted was focused on the HN and lung cohorts. For the studies focused on HN malignancies, the applications investigated included the geometric validation of deformations for multiple DIR algorithms, the assessment of the uncertainty in dose recalculation using a CBCT-based dCT, and the uncertainty in dose summation resulting from the properties of the underlying deformation fields. The dCT method was shown to be a good interim solution to repeat CT and a superior alternative to the direct use of current CBCT imaging technology for dose calculation, all in the context of ART; for proton therapy treatments the associated uncertainties of the method were in general higher than for IMRT. The ability of using DIR to co-register multimodal and multitemporal data in the HN region was also investigated; the results found were promising and the limitations of current algorithms and data acquisition protocols were identified. Following the work done on the HN cohort, a clinical workflow for lung adaptive proton therapy was proposed and evaluated, both in terms of technical and clinical parameters. This work was the first clinical evaluation of proton-gantry CBCT, and demonstrated the usefulness of image-guidance in this type of radiotherapy treatments. It is clear that there is still a lot of work to be done before

these methods and techniques become part of widespread clinical routine; however, the evidence gathered in this thesis and other publications shows that there is evidence to start translating these techniques to controlled clinical research settings. A first important point to refer is the whole problematic of DIR validation and lack of perfect gold standard to ascertain the uncertainties associated with the workflows. While throughout this thesis methodologies to bypass this issue were suggested, the lack of a true gold standard (particularly for dose summation) remains unanswered. Additionally, a lot of work has to be done to improve DIR algorithms to more accurately describe the real deformations that occur in the patient anatomy. CBCT is also a modality with its own limitations, and research effort should be done in the direction of improving its quality to become closer to CT imaging.

The use of novel artificial cancer masses as a novel platform to study the benefits of additional imaging during radiotherapy was explored. An existing artificial cancer mass model was extended to generate samples that were MRI friendly. The tumoroid model was shown to be image-able in its current state in standard T_1 and T_2 MRI acquisitions. In spite of the efforts to measure its relaxometric properties, the reproducibility and controllability of the samples were inadequate, which caused the contrast generated to be non-quantifiable. The work presented in this topic was the first exploratory work on using tissue engineering techniques for medical imaging; therefore there is a large scope of possible improvements and understanding of the current model.

IGRT and ART are topics of high relevance in clinical radiotherapy, and the work conducted in this thesis is a step toward understanding the applicability and limitations of different computational methods in clinical routine. It is clear that there is a lot of work still to be conducted before ART is implemented in a cost-efficient, accurate, optimised, automatic and widely available fashion; therefore, this is an area where research efforts are still needed.

Appendix A

Clinical indicators of replanning

Table A.1 presents the values found for under/over-ranges statistics using the dCT method and acquiring a new rCT, for all the patients included in this study. Similarly, Table A.2 shows the variation in DVH statistics from the planning values calculated using the dCT method ($D_{dCT\text{-}WET}$ and D_{dCT}) and acquiring a new rCT.

Table A.3 is a breakdown of the analysis of the clinical indicators obtained for dCT and rCT, specifying the limitations of each particular dataset, correct predictions and false negatives/positives in detections.

Table A.1: *Changes in WET between planning (pCT) and verification scans (dCT and rCT) per treatment field.*

		$WET_{under>3mm}$(%)			$WET_{over>3mm}$(%)			$WET_{under-95\%}$(mm)			$WET_{over-95\%}$(mm)		
PT#	Field	dCT	rCT	Δ	dCT	rCT	Δ	dCT	rCT	Δ	dCT	rCT	Δ
1	LPO_1	14.7	28.3	-13.6	10.4	14.3	-3.9	6.9	6.4	0.5	4.1	4.0	0.1
	LPO_2	18.9	26.6	-7.7	12.3	18.7	-6.4	34.9	16.4	18.5	19.8	15.8	3.9
2	RPO_1	0.6	0.1	0.5	0.0	0.0	0.0	50.2	54.1	-3.9	27.0	24.6	2.4
	RPO_2	0.9	0.0	0.9	0.0	0.0	0.0	57.9	61.8	-3.9	29.6	25.2	4.3
3	RPO	0.0	0.0	0.0	0.0	0.0	0.0	76.9	50.5	26.4	48.0	41.4	6.6
	PA	0.0	0.0	0.0	0.0	0.0	0.0	57.1	57.1	0.0	47.1	44.1	3.0
4	PA	0.1	0.0	0.1	0.0	0.0	0.0	33.3	34.7	-1.4	15.6	10.6	5.1
	RPO	0.0	0.0	0.0	0.0	0.0	0.0	61.8	52.5	9.3	19.1	11.0	8.1
5	PA	0.0	0.0	0.0	0.0	0.0	0.0	47.3	24.6	22.7	8.9	7.2	1.7
	LAO	0.0	0.0	0.0	0.0	0.0	0.0	17.5	15.7	1.8	8.1	5.4	2.8
6	PA	0.0	0.0	0.0	0.0	0.0	0.0	2.7	0.0	2.7	2.3	1.3	1.0
	LPO	0.0	2.8	-2.8	0.0	1.8	-1.8	6.9	0.8	6.1	3.5	1.1	2.5
7	PA	0.0	0.5	-0.5	0.0	0.0	0.0	62.8	68.9	-6.1	10.7	11.0	-0.3
	RPO	0.1	0.4	-0.2	0.0	0.0	0.0	59.2	73.3	-14.1	9.3	11.6	-2.4
8	ASO	0.3	0.0	0.3	0.0	0.0	0.0	26.8	41.9	-15.1	10.3	21.4	-11.0
	PA	0.0	0.0	0.0	0.0	0.0	0.0	60.0	41.4	18.6	15.0	25.3	-10.3
9	PA	0.0	0.0	0.0	0.0	0.0	0.0	50.3	7.2	43.1	8.8	3.7	5.1
	LPO	0.0	0.0	0.0	0.0	0.0	0.0	68.6	17.1	51.5	9.6	5.2	4.4
10	PA	0.0	1.2	-1.2	0.0	1.3	-1.3	5.0	0.0	5.0	3.0	0.0	3.0
	LPO	0.0	1.0	-0.9	0.0	0.8	-0.8	11.0	0.8	10.2	4.6	0.0	4.5

11	AP	0.0	0.0	0.0	0.0	0.0	0.0	8.5	29.2	-20.7	4.1	8.5	-4.4
	RPO	0.0	1.2	-1.1	0.0	0.0	0.0	15.0	14.7	0.3	5.8	7.5	-1.7
12	PA	0.0	0.0	0.0	0.0	0.0	0.0	19.2	39.7	-20.5	5.4	9.3	-3.9
	RPO	0.0	0.0	0.0	0.0	0.0	0.0	31.9	48.2	-16.4	8.3	10.5	-2.2
13	PA	0.2	0.0	0.2	0.0	0.0	0.0	14.8	33.8	-18.9	10.5	16.0	-5.5
	LPO	0.0	0.0	0.0	0.0	0.0	0.0	35.1	43.4	-8.3	15.5	17.6	-2.1
14	PA	0.8	0.7	0.1	0.0	0.0	0.0	39.0	42.2	-3.1	19.5	26.4	-6.9
	LPO	1.2	2.2	-1.0	0.0	0.0	0.0	42.6	37.5	5.1	22.9	30.7	-7.8
15	PA	0.0	0.1	-0.1	0.0	0.0	0.0	24.9	26.6	-1.7	6.5	9.1	-2.6
	RPO	0.0	0.3	-0.3	0.0	0.0	0.0	59.0	27.7	31.3	8.5	8.4	0.1
16	RPO	3.2	2.2	1.0	1.9	1.1	0.8	9.2	5.8	3.4	5.1	3.6	1.5
	LAO	0.7	1.9	-1.3	0.0	0.0	0.0	11.4	5.2	6.2	5.7	3.1	2.7
17	LPO_1	0.0	0.3	-0.3	0.0	0.4	-0.4	18.0	15.2	2.8	7.0	8.8	-1.8
	LPO_2	0.1	0.5	-0.4	0.0	0.3	-0.3	19.1	12.6	6.5	16.8	16.8	0.0
18	AP	1.4	0.2	1.1	0.0	0.0	0.0	2.8	1.0	1.7	2.6	1.5	1.0
	RAO	2.6	0.0	2.6	1.8	0.0	1.8	2.1	1.6	0.5	1.8	1.6	0.1
19	PA	0.0	0.0	0.0	0.0	0.0	0.0	6.1	28.7	-22.6	3.3	6.0	-2.7
	RPO	0.0	0.0	0.0	0.0	0.0	0.0	18.8	46.6	-27.8	6.6	7.6	-1.1
20	RPO_1	0.9	0.0	0.9	0.0	0.0	0.0	27.0	40.7	-13.6	17.0	13.9	3.1
	RPO_2	1.0	0.0	1.0	0.0	0.0	0.0	30.8	39.8	-9.0	17.1	13.2	3.9

Table A.2: *Variation in DVH statistics from planning to verification doses: dCT range-corrected and recalculated doses ($D_{dCT\text{-}WET}$ and D_{dCT}), and rCT recalculated dose (D_{rCT}).*

	PTV $\Delta V_{95\%}$(%)			iCTV $\Delta V_{99\%}$(%)			Heart ΔD_{max}(Gy)			Heart ΔV_{45Gy}(%)			Cord ΔD_{max}(Gy)		
PT#	$D_{dCT\text{-}WET}$	D_{dCT}	D_{rCT}	$D_{dCT\text{-}WET}$	D_{dCT}	D_{rCT}	$D_{dCT\text{-}WET}$	D_{dCT}	D_{rCT}	$D_{dCT\text{-}WET}$	D_{dCT}	D_{rCT}	$D_{dCT\text{-}WET}$	D_{dCT}	D_{rCT}
1	-3.0	-1.4	-3.1	-9.4	-10.7	-2.8	6.2	6.2	13.8	0.0	0.0	0.0	2.5	2.5	2.5
2	-0.3	-0.2	-0.6	-7.4	-6.3	-7.5	-1.4	-0.7	-0.8	5.6	5.5	9.8	6.9	6.5	3.6
3	-4.8	-5.7	-0.3	-27.4	-27.1	-13.4	0.0	-0.2	-0.4	-0.1	-0.2	-0.1	-0.2	0.0	1.3
4	0.1	0.0	-0.1	0.3	-0.6	-1.3	-0.3	-0.3	-0.1	-4.6	-4.6	-1.1	-4.2	-4.3	-0.2
5	-1.8	-1.1	-0.7	-1.4	-0.8	-0.6	-0.2	0.2	-0.1	0.5	0.8	1.0	-1.2	0.8	1.1
6	-0.2	-0.1	-0.3	0.2	0.0	0.2	3.4	5.4	9.4	0.0	0.0	0.0	3.0	3.3	1.0
7	-0.1	-0.4	-0.9	-3.6	-4.6	-7.4	-1.4	-1.1	-0.4	-0.3	-0.2	0.5	2.8	3.2	4.1
8	-1.0	-0.7	-3.2	-4.7	-3.0	-13.5	0.3	-0.1	-0.5	3.2	3.1	9.7	1.3	1.3	0.3
9	0.6	0.5	0.1	-2.0	-5.7	-0.2	-0.4	-0.5	-0.6	0.5	0.8	0.1	0.1	0.0	0.5
10	0.0	0.2	-0.6	0.4	-1.2	0.9	-0.5	-0.2	-0.1	-0.2	0.1	0.1	2.2	2.2	2.9
11	-0.2	0.0	-0.2	0.0	0.1	2.2	-0.5	0.3	-0.7	0.5	0.5	2.2	5.7	6.8	7.3
12	-0.2	-0.3	-0.1	0.2	0.0	-0.2	24.8	20.9	18.0	0.0	0.0	0.0	7.2	7.3	5.6
13	0.5	0.7	-0.1	3.1	1.4	-3.7	0.1	0.1	0.7	0.1	0.1	1.9	-2.0	-1.5	1.5
14	-0.6	-0.7	-0.8	-4.8	-7.4	-6.0	-0.9	-0.3	-0.2	0.7	0.9	1.2	-0.1	0.0	1.9
15	-0.1	0.0	-0.6	-5.8	-5.7	-1.0	-0.2	0.3	0.7	1.1	1.1	1.1	1.9	1.9	1.2
16	-0.7	-0.1	-0.1	1.4	1.4	1.9	2.9	1.5	1.8	-1.2	-0.8	0.6	10.9	11.4	4.9
17	-0.9	-0.1	-12.5	0.6	0.1	-5.0	-0.2	0.1	0.3	-0.7	-0.7	1.2	1.4	1.5	2.6
18	-0.9	-0.3	-0.1	-0.3	-0.6	0.2	-0.2	0.1	0.0	1.9	2.4	0.8	-7.4	-6.0	5.1
19	-0.2	0.1	0.0	0.6	-0.2	-1.3	1.9	2.3	0.8	0.0	0.0	0.0	1.1	1.2	1.5
20	-1.3	-1.2	-1.4	0.1	-2.6	-3.1	-0.8	1.4	-0.3	0.5	0.9	0.9	11.6	11.6	3.2

Table A.3: *Summary of the results obtained for clinical indicators extracted per patient for dCT and rCT: properties/limitations of the datasets, and analysis of the clinical indicators (correct predictions and false positives/negatives).*

PT #	corrected dCT (Y/N)	Properties and limitations of the dataset[a]	Analysis of clinical indicators Correct detections	 False negatives/positives
1	Y	-dCT corrected for atelectasis; -Differences in positioning between rCBCT and rCT (scapula)	-Significant under-ranging: $WET_{under\text{-}95\%}$ >10 mm (both fields) -Significant over-ranging: $WET_{over\text{-}95\%}$ >15 mm (LPO_2 field) -2D WET difference maps: consistent[b] -Dose to oesophagus: D_{max} from 50 Gy to 71/71/68 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$ -DVHs: loss of coverage and increase dose to oesophagus	*-Overestimation of loss of iCTV coverage: $\Delta V_{99\%}$=-9/-11/-3% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$*[c]
2	Y	-dCT corrected for tumour erosion -Differences in positioning between rCBCT and rCT (deformation of external contours)	-Significant over-ranging: $WET_{over\text{-}95\%}$ >15 mm (both fields) -2D WET difference maps: consistent -iCTV coverage: $\Delta V_{99\%}$=-7/-6/-8% at $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$) -Dose to cord: D_{max} from 45 to 52/52/49 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$) -Dose to heart: V_{45Gy} from 25% to 31/31/35% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$) -DVHs: loss of coverage and increase dose to heart/cord	

3	Y	-dCT corrected for lung reinflation -Cropped contours (RPO field)	-Significant over-ranging: $WET_{over-95\%}$ >15 mm (both fields) -2D WET difference maps: consistent -iCTV coverage: $\Delta V_{99\%}$=-27/-27/-13% for $D_{dCT-WET}/D_{dCT}/D_{rCT}$	-DVHs: overestimation of loss of coverage
4	N	-Cropped contours (RPO field)	-2D WET difference maps: consistent -DVHs: no major changes	*-Overestimation of over-ranging: $WET_{over-95\%}$=15.6/10.6 (PA field) and 19.1/11.0 (RPO field) for dCT/rCT*[d]
5	N		-2D WET difference maps: consistent -DVHs: no major changes	
6	N		-DVHs: no major changes	-2D WET difference maps: inconsistent
7	N	-Differences in positioning between rCBCT and rCT (positioning of pacemaker wires)	-iCTV coverage: $\Delta V_{99\%}$=-4/-5/-7% for $D_{dCT-WET}/D_{dCT}/D_{rCT}$ -DVHs: loss of iCTV/PTV coverage	*-Overestimation of loss of iCTV coverage: $\Delta V_{99\%}$=-27/-27/-15% for $D_{dCT-WET}/D_{dCT}/D_{rCT}$* -2D WET difference maps: inconsistent
8	Y	-dCT corrected for tumour shrinkage -Poor image quality (large patient)	-2D WET difference maps: consistent -DVHs: loss of iCTV/PTV coverage	-Overestimation of over-ranging: $WET_{over-95\%}$=10.3/21.4 (ASO field) and 15.0/25.3 (PA field) for dCT/rCT *-Underestimation of loss of iCTV coverage: $\Delta V_{99\%}$=-5/-3/-14% for $D_{dCT-WET}/D_{dCT}/D_{rCT}$* -Underestimation of dose to heart: V_{45Gy} from 22% to 25/25/31% for $D_{dCT-WET}/D_{dCT}/D_{rCT}$) -DVHs: underestimation of dose to heart
9	N	-Differences in positioning between rCBCT and rCT (scapula)	-2D WET difference maps: consistent (LPO field)	-2D WET difference maps: inconsistent (PA field) -Overestimation of loss of iCTV coverage: $\Delta V_{99\%}$=-2/-6/0% for $D_{dCT-WET}/D_{dCT}/D_{rCT}$ -DVHs: Overestimation of loss of iCTV coverage

10	N	-Differences in positioning between rCBCT and rCT (deformation of external contours)	-Differences in positioning between rCBCT and rCT (deformation of external contours)	-2D WET difference maps: consistent -DVHs: no major changes
11	N		-2D WET difference maps: consistent -Dose to cord: D_{max} from 30 to 36/37/37 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$ -DVHs: no major changes	
12	N	-Differences in positioning between rCBCT and rCT (misalignment external contours)	-2D WET difference maps: consistent (PA field) -Dose to heart: D_{max} from 1 to 26/22/19 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$ -DVHs: no major changes	-2D WET difference maps: inconsistent (RPO field)
13	N	-Differences in positioning between rCBCT and rCT (positioning of airways)	-2D WET difference maps: consistent -Significant over-ranging: $WET_{over\text{-}95\%}$ >15 mm (LPO fields) -DVHs: no major changes	-Underestimation of over-ranging: $WET_{over\text{-}95\%}$=10.5/16.0 (PA field) -Different predictions of tumours coverage ($\Delta V_{99\%}$=+3/+1/-4% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$).
14	Y	-dCT corrected for tumour shrinkage	-Significant over-ranging: $WET_{over\text{-}95\%}$ >15 mm (both fields) -2D WET maps: consistent -iCTV coverage: $\Delta V_{99\%}$=-5/-7/-6% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$ -DVHs: loss of iCTV/PTV coverage	
15	N	-Cropped contours (RPO field) -Differences in positioning between rCBCT and rCT (misalignment external contours)	-Dose to heart: V_{30Gy} from 10 to 14/14/14% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$ -DVHs: increase in heart dose	-2D WET difference maps: inconsistent -Overestimation of loss of iCTV coverage: $\Delta V_{99\%}$=-6/-6/-1% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$

16	N	-Differences in positioning between rCBCT and rCT	-2D WET difference maps: consistent -Dose to cord: D_{max} from 35 to 46/47/40 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$ -DVHs: right shift of iCTV/PTV DVH curves	
17	N	-Cropped contours (LPO field)	-2D WET difference maps: consistent -Significant over-ranging: $WET_{over\text{-}95\%}$ >15 mm (LPO_2 field) -DVHs: no major changes	-Different predictions of tumours coverage iCTV $\Delta V_{99\%}$ =0/0/-5%)
18	N		-2D WET difference maps: consistent -DVHs: no major changes	
19	N	-Differences in positioning between rCBCT and rCT (scapula)	-DVHs: no major changes	-2D WET difference maps: inconsistent -Underestimation of over-ranging: WETover>3mm was 18.8/46.6% for dCT/rCT (RPO field)
20	Y	-dCT corrected for tumour shrinkage -Change in tumour density	-2D WET maps: consistent -DVHs: target coverage loss	-Underestimation of loss of iCTV coverage: $\Delta V_{99\%}$=0/-3/-3% for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$ -Overestimation of dose to cord: D_{max} from 27 to 39/39/30 Gy for $D_{dCT\text{-}WET}/D_{dCT}/D_{rCT}$

[a] Only differences in setup that are in beam path are described.

[b] Consistency in 2D WET maps refers to similar topology of under/over-range location.

[c] Differences in $\Delta V_{99\%}$ are reported as false positives/negatives; however if the DVH curves did not show the same behaviour this was not considered a failure of the method. $\Delta V_{99\%}$ was found to be a sensitive parameter. Such cases were highlighted in italic font.

[d] Large discrepancy in $WET_{over\text{-}3mm}/WET_{under\text{-}3mm}$ were reported as false positives/negative; however if the 2D WET maps were consistent this was not considered a failure of the method. $WET_{over\text{-}3mm}/WET_{under\text{-}3mm}$ were found to be quite sensitive parameters. Such cases are were highlighted in italic font.

Appendix B

Cell maintenance protocol

Material

- pipettes (5 and 10 mL)
- 75cm^2 flask(s)

Reagents

- DMEM + 10% FBS + 1% P/S
- phosphate buffed saline (PBS)

Protocol

1. Warm up the reagents in a water bath (≈37°).
2. Prepare all the material needed in the hood.
3. Remove the cells from the incubator and view cultures with a microscope to assess the degree of confluency and confirm the absence of bacterial and fungal contaminants.
4. Pipette the entire medium from the flask.
5. Add ≈12 mL of fresh medium to the flask.
6. Ensure all cell surface is covered by the medium, and return the sample to the incubator.

Appendix C

Cell subculture protocol

Material

- pipettes (5 and 10 mL)
- 75cm^2 flask(s)

Reagents

- DMEM + 10% FBS + 1% P/S
- 1× Trypsen-ethylenediaminetetraacetic acid (EDTA)
- PBS

Protocol

1. Warm up the reagents in a water bath (≈37°).
2. Prepare all the material needed in the hood.
3. Remove the cells from the incubator and view cultures with a microscope to assess the degree of confluency and confirm the absence of bacterial and fungal contaminants.
4. Pipette the entire medium from the flask.
5. Wash the cells with 10 mL of saline solution, PBS. Do not pour directly on the cell surface. PBS washing removes any remaining medium and associated proteins that can neutralise the detaching enzyme. Repeat if there are any vestiges of the medium left.

6. Add 3 mL of Trypsin-EDTA, a detaching enzyme. Although most cells will detach in the presence of trypsin alone the EDTA is added to enhance the activity of the enzyme. Ensure the whole surface is covered.
7. Leave in the incubator for 5 minutes.
8. Meanwhile, add 12 mL of medium to the flask(s) the cells are going to be passaged to.
9. Examine the cells using an inverted microscope to ensure that all the cells are detached and floating. The side of the flasks may be gently tapped to release any remaining attached cells.
10. Add 7 mL of medium and gently mix with the pipette. The medium neutralises the Trypsin, that when in contact with the cells for too long starts breaking the membranes and killing them. The ratio of Trypsin:medium should always be at least 2:1.
11. Transfer a fraction of the final solution (10 mL) to the new flask(s). At least 10% of the original detached population must be passaged.
12. Store the flasks in the incubator, ensuring that the solution is homogeneously distributed in the surface.

Appendix D

Cell counting protocol

Material

- haemocytometer
- pipettes (5 and 10 mL)
- Gilson pipettes (P20, P1000)
- universal tube (20 mL)

Reagents

- cell suspension

Protocol

1. Follow steps 1-11 of Protocol C to prepare the cell suspension.
2. Transfer the cell solution into a tube, and centrifuge (1500 rpm for 5 minutes). Do not forget to balance the weight with a tube filled with approximately the same volume of water.
3. Pipette the entire medium from the tube. Avoid disturbing the cell pellet by tilting the tube slightly while removing all the media.
4. Re-suspend the cells in a fixed volume of medium (V=4-6 mL, depending on confluence of the flask initially). Gently mix with the pipette, ensure the solution is homogeneous.
5. Remove 10 μL from the solution and flood the haemocytometer chamber using a Gilson pipette.

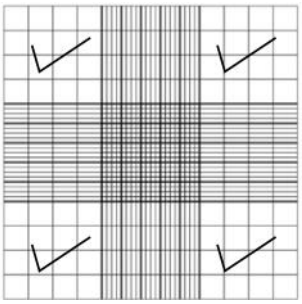

Figure D.1: *Squares of the haemocytometer used in cell counting.*

6. Count in the microscope the number of cells in 4 of the 9 of the large squares of the grid (Figure D.1). Average the number of cells counted (N_{count}).

7. The total number of cells in suspension is given by:

$$N_{cells} = N_{count} \times C_{haem} \times V \tag{D.1}$$

C_{haem} is the conversion factor of the haemocytometer, and is the inverse of the volume of each of the large squares. For the haemocytometer used in this work it corresponded to a conversion factor of 10^4.
Example: If N_{count}=250 and V=5 mL, then the solution contains a total of:

$$N_{cells} = 250 \times 10^4 \times 5 = 12.5 \times 10^6 \text{ cells} \tag{D.2}$$

A dilution factor of the solution may also have to be considered if the cell suspension was diluted in other volume, such as Trypan blue. Trypan blue is a vital stain used to selectively colour dead tissues or cells blue, and is commonly used to check the viability of the cells.

8. Centrifuge again the cell solution (1500 rpm for 5 minutes).

9. Pipette the entire medium from the tube.

10. Re-suspend the cells in the desired volume of medium ($V_{suspension}$). For tumoroid preparation, the amount of cells desired must be diluted in 0.4 mL of medium. Considering that we want to include a total of $N_{tumoroid}$ cells in the collagen gel, the volume to re-suspend the cells in ($V_{solution}$) is calculated by:

$$V_{solution} = \frac{0.4\,\text{mL}}{N_{tumoroid}} \times N_{cells} \tag{D.3}$$

Example: To produce a 4 mL gel with 6.4×10^6 cells using the previous example solution (N_{cells}=12.5×10^6), the cells have to be re-suspended in:

$$V_{solution} = \frac{0.4\,\text{mL}}{N_{tumoroid}} \times N_{cells} = \frac{0.4\,\text{mL}}{6.4 \times 10^6} \times 12.5 \times 10^6 = 0.78 \text{ mL} \tag{D.4}$$

Note: HT29 cells by gravity action start to deposit at the bottom of the container with time. Carefully re-suspend the solution (mix gently) before using it in any step of the process. This ensures a homogeneous distribution of cells in solution.

Appendix E

Collagen matrix preparation protocol

Material

- universal tube (20 mL)
- Petri dish
- ice container
- syringes (1 mL)
- needles (21 g)
- pipettes (2 mL and 5 mL)
- pipette tips (20 μL and 1 mL)
- Pasteur pipette
- autoclaved bag: mould and plunger, filter paper, glass slide, nylon meshes, steel meshes, tweezers
- surgical scalpel

Reagents

- 10× MEM
- DMEM + 10% FBS + 1% P/S
- 1M and 5M NaOH (filtered and sterile – do not use if prepared more than 6 months ago)
- cell suspension (Appendix D)
- collagen Type I

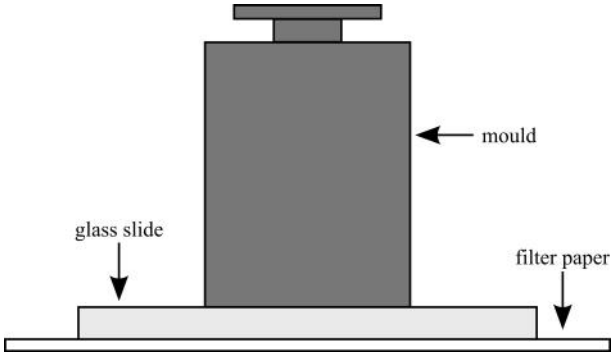

Figure E.1: *Mould preparation for tumoroid production: filter paper, glass slide and mould. The plunger should not be inside the mould at this point (it is in the figure just to be indicative of the orientation of the mould). All the equipment must be autoclaved.*

Protocol

1. Prepare all the materials in the hood. Store the collagen and the universal tube inside an ice container.
2. Prepare the syringes for the pH neutralisation. Add a small volume (≈0.2 mL) of 5M NaOH to one, and a larger volume (≈0.5 mL) of 1M NaOH to the other. Properly identify each of the syringes by its content.
3. Prepare the mould (Figure E.1).
4. Pipette 0.4 mL of 10× MEM into the universal tube.
5. Pipette 3.2 mL of collagen type I solution into the universal tube. When handling the collagen it is important to not introduce any air bubbles. Pour the collagen gently against the tube wall, and mix the tube gently to ensure a homogenous yellow colour throughout.
6. Neutralise the solution by adding 5M NaOH (filtered) dropwise into the collagen-MEM solution until the colour changes from yellow to pink, and back to a light orange. Keep the tube inside the ice container as much as possible.
7. Neutralise the solution by adding 1M NaOH dropwise into the collagen-MEM solution until the colour changes from yellow to bright pink. Take care not to over-neutralise the solution as the gel will not set otherwise.
8. Pipette 0.4 mL of cell suspension into the neutralised collagen solution. Mix gently with the pipette tip, and by gently shaking the tube.
9. At this stage, take care that the collagen gel has already begun to set. Transfer the 4ml solution into the mould using the Pasteur pipette. To avoid bubbles do not overfill the pipette and/or dispose of all its contents in one go. Always leave the last drop on the pipette to check the gel setting.

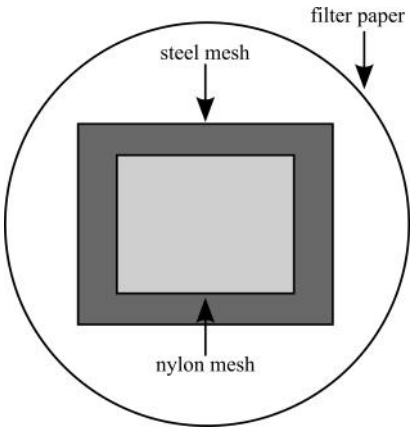

Figure E.2: *Preparation of the meshes for tumoroid production: filter paper, steel mesh and nylon mesh. All the equipment must be autoclaved.*

10. Allow the collagen solution to become a gel by leaving for 30 minutes at room temperature.
11. Prepare the material for plastic compression (Figure E.2). Place the steel mesh on top of the filter paper and place the nylon mesh above the steel mesh. Always hold the meshes in the corner to avoid contamination.
12. After the 30 minutes, place the meshes arrangement on top of the mould, and flip over. From the side gently lift the glass slide, and hold the apparatus as a with both hands from the sides. Only touch the meshes in the corner to avoid contamination. Place the weight of the plunger (175g) within its slot for 30 seconds to compress the gel.
13. Repeat compression on the other side of the gel.
14. Cut up the gel inside the Petri dish into 2 equal pieces (by eye). Place media on top of the gels and store in the incubator.
15. Follow steps 4 to 8 to create an acellular hydrogels. Use 0.4 mL of DMEM instead of cell suspension in the solution. Divide the acellular mixture into 2 containers, and then place a piece of the ACM) within the hydrogel. Return to the incubator.
16. After the gels sets (≈20 minutes), place 1 mL of media directly on top and return to the incubator. The media should be changed every 2 days.

Appendix F

Mould re-design

The possibility to re-design the mould used to compress the gel was investigated for two main reasons. First, the plastic compression process had to be optimised to achieve the highest density, and the effect of the varying surface in absorbing water out was important to be understood. Secondly, the samples produced with the old mould had anisotropic dimensions, and a more suitable geometry was desired to maximise imaging contrast due to system resolution. The aim of this experiment was to (i) evaluate if differences in surface area negatively affect the plastic compression process and (ii) measure the final volume and collagen density of the samples.

Different mould re-design strategies were considered (Figure F.1). All of the proposed designs required new plunger(s), and had its own pros and cons.

1. New mould with two holes of varying surface. The main advantage of this design is its simplicity and similarity with the old mould. It requires a new mould to be made for the purpose of testing.

2. Modify old mould by adding a metal filler. The advantage of this method is to use something already available, and just modify it to meet the testing needs. The cost associated with building is therefore smaller. The simplest way to attach the filler to the main body is using autoclave tape. However, several deficiencies occur with this design. First, inserts have to be mounted and adhered to the main body before autoclaving, adding risk to metal expansion and consequent difficulty in removing the inserts later (the current mould can be split in two parts, which makes this a solvable issue). The fillers would have to be made of heavy stainless steel, and tape may not be strong enough to hold everything in place, making difficult the handling of the device during plastic compression.

3. Design a mould with varying surface. The main advantage of this design is its versatility; it could be used for a wide range of applications apart from imaging. Therefore, it is more suitable as an end-point hardware, instead of a test equipment. The additional material necessary will significantly increase its weight. Ergonomy

is also reduced, as flipping might be difficult with the extras. To reduce problems with autoclaving and expansion, the mould could open and close. Screws would allow the system to be stable in any position chosen.

A simple experiment was designed to assess the usefulness of the mould re-design. Moulds of different surface areas were used to prepare acellular collagen gels of different dimensions. Since the main aim of this experiment was to check if the surface area will affect the effectiveness of the plastic compression, and therefore exact optimum radius cannot be predicted without knowing the compression achievable, design (a) was manufactured since it was easy and fast to produce. It was built in aluminium because of the easiness to handle this metal and it contained 2 circular holes with different radius (8 and 12 mm). A circular shape was chosen to make the manufacture easier, since holes could be easily drilled in the bulk material (Figure F.2). The values chosen for the hole size were such that the surface area was approximately halved (and therefore height doubled).

This mould could not be autoclaved and was only used to produce non-aseptic samples. Two acellular tumoroids were prepared and compressed using the two holes of the re-designed mould. The values of density measured (ratio between dry and wet weight) were 2.19% and 1.91% for the larger and smaller surfaces, respectively. This provides evidence that reducing the surface area impacted the final density of the construct, and therefore in the imaging experiments described in this thesis the previously available mould with additional load was used.

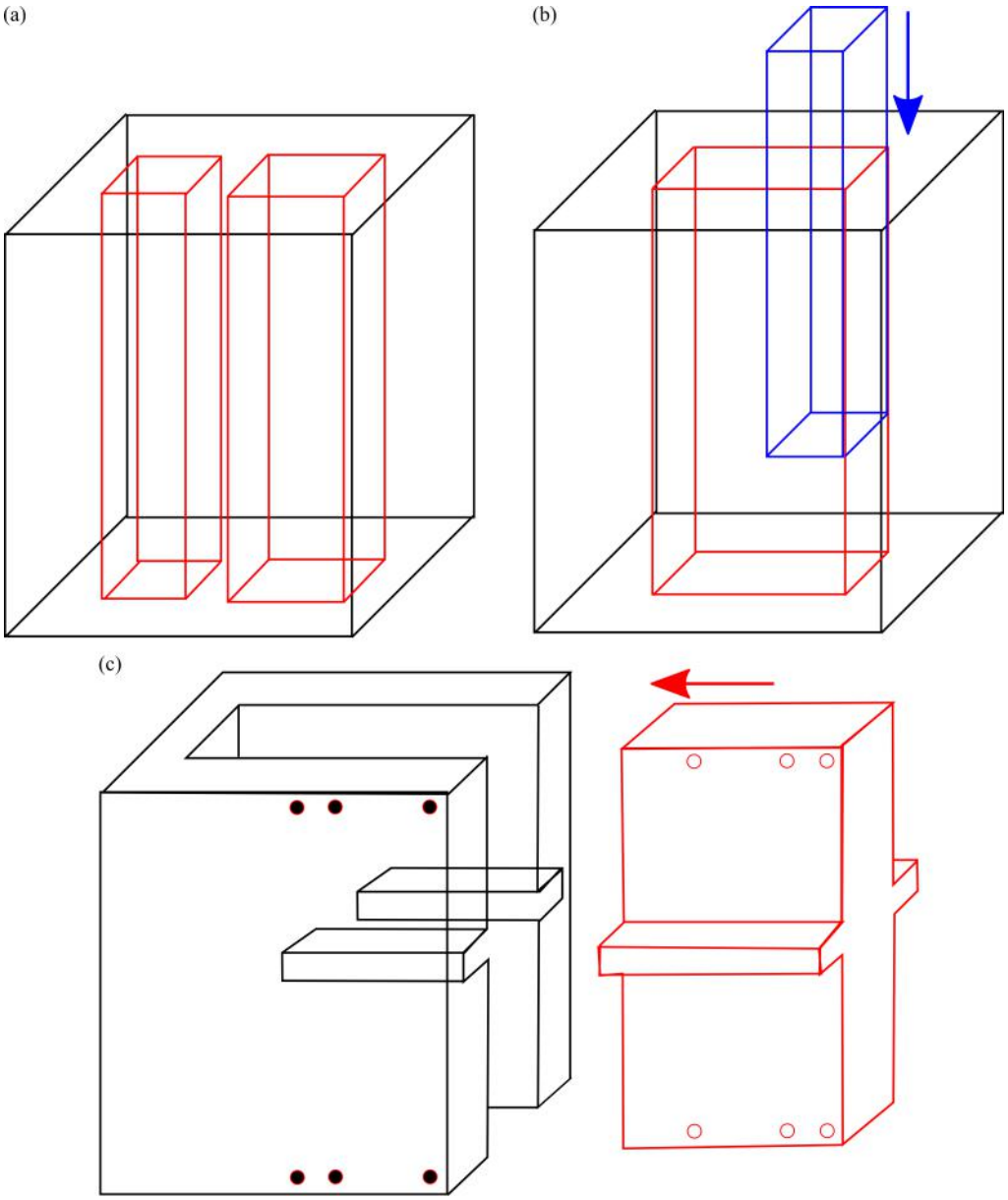

Figure F.1: *Mould re-design suggestions: (a) new mould with two holes of varying surface; (b) modifying old mould by adding a metal filler; and (c) new mould with varying surface.*

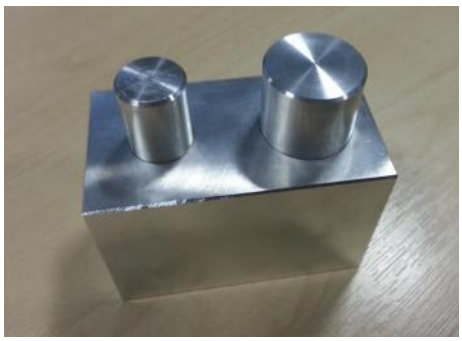

Figure F.2: *Re-designed mould. Built in aluminium, it consisted of two circular holes with different radius (8 and 12 mm).*

Bibliography

[1]E. J. Hall, *Radiobiology for the radiologist* (Lippincott Williams & Wilkins, 2006) (cit. on pp. 27, 38–40).

[2]S. Webb, "The physical basis of IMRT and inverse planning", The British Journal of Radiology **76**, 678–689 (2003) (cit. on p. 27).

[3]T. Bortfeld, "IMRT: a review and preview", Physics in Medicine and Biology **51**, R363–379 (2006) (cit. on p. 27).

[4]R. R. Wilson, "Radiological use of fast protons", Radiology **47**, 487–491 (1946) (cit. on pp. 27, 101).

[5]S. A. Bhide and C. M. Nutting, "Recent advances in radiotherapy", BMC Medicine **8**, 1–5 (2010) (cit. on p. 28).

[6]D. Yang, S. Brame, I. El Naqa, A. Aditya, Y. Wu, S. M. Goddu, S. Mutic, J. O. Deasy, and D. A. Low, "Technical note: DIRART–a software suite for deformable image registration and adaptive radiotherapy research", Medical Physics **38**, 67–77 (2011) (cit. on p. 28).

[7]O. A. Zeidan, K. M. Langen, S. L. Meeks, R. R. Manon, T. H. Wagner, T. R. Willoughby, D. W. Jenkins, and P. A. Kupelian, "Evaluation of image-guidance protocols in the treatment of head and neck cancers", International Journal of Radiation Oncology, Biology, Physics **67**, 670–677 (2007) (cit. on p. 28).

[8]D. Yan, F. Vicini, J. Wong, and A. Martinez, "Adaptive radiation therapy", Physics in Medicine and Biology **42**, 123–132 (1997) (cit. on pp. 28, 50).

[9]Y. Yang, E. Schreibmann, T. Li, C. Wang, and L. Xing, "Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation", Physics in Medicine and Biology **52**, 685–705 (2007) (cit. on pp. 31, 50, 68, 70, 95, 114, 150).

[10]M. Peroni, D. Ciardo, M. F. Spadea, M. Riboldi, S. Comi, D. Alterio, G. Baroni, and R. Orecchia, "Automatic segmentation and online virtualCT in head-and-neck adaptive radiation therapy", International Journal of Radiation Oncology, Biology, Physics **84**, e427–433 (2012) (cit. on pp. 31, 50, 68, 82, 86, 150).

[11]G. Landry, G. Dedes, C. Zöllner, J. Handrack, G. Janssens, J. O. d. Xivry, M. Reiner, C. Paganelli, M. Riboldi, F. Kamp, M. Söhn, J. J. Wilkens, G. Baroni, C. Belka, and K. Parodi, "Phantom based evaluation of CT to CBCT image registration for proton therapy dose recalculation", Physics in Medicine and Biology **60**, 595 (2015) (cit. on pp. 31, 50, 114, 118, 128, 150).

[12]G. Landry, R. Nijhuis, G. Dedes, J. Handrack, C. Thieke, G. Janssens, J. Orban de Xivry, M. Reiner, F. Kamp, J. J. Wilkens, C. Paganelli, M. Riboldi, G. Baroni, U. Ganswindt, C. Belka, et al., "Investigating CT to CBCT image registration for head and neck proton therapy as a tool for daily dose recalculation", Medical Physics **42**, 1354–1366 (2015) (cit. on pp. 31, 114, 115, 118, 128, 150).

[13]V. L. Rosa, A. Kacperek, G. Royle, and A. Gibson, "Range verification for eye proton therapy based on proton-induced x-ray emissions from implanted metal markers", Physics in Medicine and Biology **59**, 2623 (2014) (cit. on p. 33).

[14]A. K. Hoang Duc, "Atlas-based methods in radiotherapy treatment of head and neck cancer", PhD thesis (University College London, 2015) (cit. on p. 33).

[15]B. S. C. Study, "Understanding cancer", (2007) (cit. on p. 38).

[16]M. C. Joiner and A. v. d. Kogel, "Basic clinical radiobiology", (2009) (cit. on p. 38).

[17]M. Najafi, R. Fardid, G. Hadadi, and M. Fardid, "The mechanisms of radiation-induced bystander effect", Journal of Biomedical Physics & Engineering **4**, 163–172 (2014) (cit. on p. 39).

[18]P. Mayles, A. Nahum, and J. C. Rosenwald, *Handbook of radiotherapy physics: theory and practice* (CRC Press, New York, 2007), 1450 pp. (cit. on pp. 39, 40, 61).

[19]K. N. Prasad, "Handbook of radiobiology", (1995) (cit. on p. 39).

[20]M. Thomas, *Biomedical image processing* (Springer-Verlag Berlin Heidelberg, 2011) (cit. on pp. 41, 46).

[21]D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images", IEEE Transactions on Medical Imaging **18**, 712–721 (1999) (cit. on pp. 42, 44, 46, 69, 87).

[22]M. Ferrant, B. Macq, A. Nabavi, and S. K. Warfield, "Deformable modeling for characterizing biomedical shape changes", Discrete Geometry for Computer Imagery, 235–248 (2000) (cit. on p. 42).

[23]G. E. Christensen, R. D. Rabbit, M. I. Miller, S. C. Joshi, U. Grenander, and T. Coogan, "Topological properties of smooth anatomic maps", Information Processing in Medical Imaging, 101–112 (1995) (cit. on p. 42).

[24]J. P. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons", Medical Image Analysis **2**, 243–260 (1998) (cit. on p. 42).

[25]S. Lee, G. Wolberg, K.-y. Chwa, and S. Y. Shin, "Image metamorphosis with scattered feature constraints", IEEE Transactions on Visualization and Computer Graphics **2**, 337–354 (1996) (cit. on p. 42).

[26]T. W. Sederberg and S. R. Parry, "Free-form deformation of solid geometric models", Computer Graphics **20**, 151–160 (1986) (cit. on p. 42).

[27]S. Lee, G. Wolberg, and S. Y. Shin, "Scattered data interpolation with multilevel B-splines", IEEE Transactions on Visualization and Computer Graphics **3**, 228–244 (1997) (cit. on p. 42).

[28]W. M. Hsu, J. F. Hughes, and H. Kaufman, "Direct manipulation of free-form deformations", 177–184 (1992) (cit. on p. 42).

[29]W. Birkfellner, M. Stock, M. Figl, C. Gendrin, J. Hummel, S. Dong, J. Kettenbach, D. Georg, and H. Bergmann, "Stochastic rank correlation: a robust merit function for 2D/3D registration of image data obtained at different energies", Medical Physics **36**, 3420–3428 (2009) (cit. on pp. 43, 82).

[30]D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers, "Diffeomorphic registration using B-splines", Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (Copenhagen, Denmark, 2006), 702–709 (2006) (cit. on p. 45).

[31]G. Christensen and H. J. Johnson, "Consistent image registration", IEEE Transactions on Medical Imaging **20**, 568–582 (2001) (cit. on p. 45).

[32]B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain", Medical Image Analysis **12**, 26–41 (2008) (cit. on p. 46).

[33]V. Arsigny, O. Commowick, X. Pennec, and N. Ayache, "A log-euclidean framework for statistics on diffeomorphisms", Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (Copenhagen, Denmark, 2006) **9**, 924–931 (2006) (cit. on p. 46).

[34]S. Ourselin, A. Roche, G. Subsol, X. Pennec, and N. Ayache, "Reconstructing a 3D structure from serial histological sections", Image and Vision Computing **19**, 25–31 (2001) (cit. on p. 46).

[35]M. Modat, G. R. Ridgway, Z. A. Taylor, M. Lehmann, J. Barnes, D. J. Hawkes, N. C. Fox, and S. Ourselin, "Fast free-form deformation using graphics processing units", Computer Methods and Programs in Biomedicine **98**, 278–284 (2010) (cit. on p. 46).

[36]M. Modat, M. J. Cardoso, P. Daga, D. Cash, N. C. Fox, and S. Ourselin, "Inverse-consistent symmetric free form deformation", Biomedical Image Registration, 79–88 (2012) (cit. on pp. 46, 69, 87).

[37]M. Modat, P. Daga, M. Cardoso, S. Ourselin, G. Ridgway, and J. Ashburner, "Parametric non-rigid registration using a stationary velocity field", Proceeding of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (Breckenridge, CO, USA, 2012), 145–150 (2012) (cit. on pp. 47, 89, 107).

[38]C. Yan, H. Zhong, M. Murphy, E. Weiss, and J. V. Siebers, "A pseudoinverse deformation vector field generator and its applications", Medical Physics **37**, 1117–1128 (2010) (cit. on pp. 47, 99).

[39]E. T. Bender, N. Hardcastle, and W. A. Tomé, "On the dosimetric effect and reduction of inverse consistency and transitivity errors in deformable image registration for dose accumulation", Medical Physics **39**, 272–280 (2012) (cit. on pp. 47, 86, 99).

[40]O. S. Pianykh, *Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide*, 2nd edition (Springer, Heidelberg ; New York, 2011), 420 pp. (cit. on p. 47).

[41]G. Janssens, L. Jacques, J. Orban de Xivry, X. Geets, B. Macq, G. Janssens, L. Jacques, J. Orban de Xivry, X. Geets, and B. Macq, "Diffeomorphic registration of images with

variable contrast enhancement, diffeomorphic registration of images with variable contrast enhancement", International Journal of Biomedical Imaging **2011**, e891585 (2010) (cit. on pp. 49, 123).

[42]J. Chen, O. Morin, M. Aubin, M. K. Bucci, C. F. Chuang, and J. Pouliot, "Dose-guided radiation therapy with megavoltage cone-beam CT", The British Journal of Radiology **79**, S87–98 (2006) (cit. on p. 49).

[43]J. L. Barker, Jr, A. S. Garden, K. K. Ang, J. C. O'Daniel, H. Wang, L. E. Court, W. H. Morrison, D. I. Rosenthal, K. S. C. Chao, S. L. Tucker, R. Mohan, and L. Dong, "Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system", International Journal of Radiation Oncology, Biology, Physics **59**, 960–970 (2004) (cit. on pp. 49, 50, 70).

[44]M. Ghilezan, D. Yan, J. Liang, D. Jaffray, J. Wong, and A. Martinez, "Online image-guided intensity-modulated radiotherapy for prostate cancer: how much improvement can we expect? A theoretical assessment of clinical benefits and potential dose escalation by improving precision and accuracy of radiation delivery", International Journal of Radiation Oncology, Biology, Physics **60**, 1602–1610 (2004) (cit. on p. 49).

[45]J. R. McClelland, S. Hughes, M. Modat, A. Qureshi, S. Ahmad, D. B. Landau, S. Ourselin, and D. J. Hawkes, "Inter-fraction variations in respiratory motion models", Physics in Medicine and Biology **56**, 251–272 (2011) (cit. on p. 50).

[46]E. K. Hansen, M. K. Bucci, J. M. Quivey, V. Weinberg, and P. Xia, "Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer", International Journal of Radiation Oncology, Biology, Physics **64**, 355–362 (2006) (cit. on p. 50).

[47]Q. J. Wu, T. Li, Q. Wu, and F.-F. Yin, "Adaptive radiation therapy: technical components and clinical applications", Cancer Journal **17**, 182–189 (2011) (cit. on pp. 50, 52).

[48]L. Xing, J. Siebers, and P. Keall, "Computational challenges for image-guided radiation therapy: framework and current research", Seminars in Radiation Oncology **17**, 245–257 (2007) (cit. on p. 50).

[49]M. Stock, M. Pasler, W. Birkfellner, P. Homolka, R. Poetter, and D. Georg, "Image quality and stability of image-guided radiotherapy (IGRT) devices: a comparative study", Radiotherapy and Oncology **93**, 1–7 (2009) (cit. on pp. 50, 81).

[50]I. Fotina, J. Hopfgartner, M. Stock, T. Steininger, C. Lütgendorf-Caucig, and D. Georg, "Feasibility of CBCT-based dose calculation: comparative analysis of HU adjustment techniques", Radiotherapy and Oncology **104**, 249–256 (2012) (cit. on p. 50).

[51]S. Y. Tsuji, A. Hwang, V. Weinberg, S. S. Yom, J. M. Quivey, and P. Xia, "Dosimetric evaluation of automatic segmentation for adaptive IMRT for head-and-neck cancer", International Journal of Radiation Oncology, Biology, Physics **77**, 707–714 (2010) (cit. on pp. 50, 60, 68, 82).

[52]P. Castadot, J. A. Lee, X. Geets, and V. Grégoire, "Adaptive radiotherapy of head and neck cancer", Seminars in Radiation Oncology **20**, 84–93 (2010) (cit. on p. 50).

[53]D. L. Schwartz and L. Dong, "Adaptive radiation therapy for head and neck cancer-can an old goal evolve into a new standard?", Journal of Oncology **2011** (2011) (cit. on p. 50).

[54]T. Zhang, Y. Chi, E. Meldolesi, and D. Yan, "Automatic delineation of on-line head-and-neck computed tomography images: toward on-line adaptive radiotherapy", International Journal of Radiation Oncology, Biology, Physics **68**, 522–530 (2007) (cit. on pp. 50, 60, 68, 82, 86).

[55]W. Lu, G. H. Olivera, Q. Chen, K. J. Ruchala, J. Haimerl, S. L. Meeks, K. M. Langen, and P. A. Kupelian, "Deformable registration of the planning image (kVCT) and the daily images (MVCT) for adaptive radiation therapy", Physics in Medicine and Biology **51**, 4357–4374 (2006) (cit. on pp. 50, 55, 82, 86, 88, 114).

[56]Q. Wu, Y. Chi, P. Y. Chen, D. J. Krauss, D. Yan, and A. Martinez, "Adaptive replanning strategies accounting for shrinkage in head and neck IMRT", International Journal of Radiation Oncology, Biology, Physics **75**, 924–932 (2009) (cit. on pp. 50, 86, 88).

[57]B. L. T. Ramaekers, M. Pijls-Johannesma, M. A. Joore, P. van den Ende, J. A. Langendijk, P. Lambin, A. G. H. Kessels, and J. P. C. Grutters, "Systematic review and meta-analysis of radiotherapy in various head and neck cancers: comparing photons, carbon-ions and protons", Cancer Treatment Reviews **37**, 185–201 (2011) (cit. on pp. 50, 104).

[58]J. Ferlay, P. Autier, M. Boniol, M. Heanue, M. Colombet, and P. Boyle, "Estimates of the cancer incidence and mortality in europe in 2006", Annals of Oncology **18**, 581–592 (2007) (cit. on p. 50).

[59]X. Geets, M. Tomsej, J. A. Lee, T. Duprez, E. Coche, G. Cosnard, M. Lonneux, and V. Grégoire, "Adaptive biological image-guided IMRT with anatomic and functional imaging in pharyngo-laryngeal tumors: impact on target volume delineation and dose distribution using helical tomotherapy", Radiotherapy and Oncology **85**, 105–115 (2007) (cit. on pp. 50, 156, 170).

[60]E. M. Vásquez Osorio, M. S. Hoogeman, A. Al-Mamgani, D. N. Teguh, P. C. Levendag, and B. J. M. Heijmen, "Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration", International Journal of Radiation Oncology, Biology, Physics **70**, 875–882 (2008) (cit. on p. 50).

[61]C. Lee, K. M. Langen, W. Lu, J. Haimerl, E. Schnarr, K. J. Ruchala, G. H. Olivera, S. L. Meeks, P. A. Kupelian, T. D. Shellenberger, and R. R. Mañon, "Evaluation of geometric changes of parotid glands during head and neck cancer radiotherapy using daily MVCT and automatic deformable registration", Radiotherapy and Oncology **89**, 81–88 (2008) (cit. on pp. 50, 70).

[62]C. Lee, K. M. Langen, W. Lu, J. Haimerl, E. Schnarr, K. J. Ruchala, G. H. Olivera, S. L. Meeks, P. A. Kupelian, T. D. Shellenberger, and R. R. Mañon, "Assessment of parotid gland dose changes during head and neck cancer radiotherapy using daily megavoltage computed tomography and deformable image registration", International Journal of Radiation Oncology, Biology, Physics **71**, 1563–1571 (2008) (cit. on p. 50).

[63]J. C. O'Daniel, A. S. Garden, D. L. Schwartz, H. Wang, K. K. Ang, A. Ahamad, D. I. Rosenthal, W. H. Morrison, J. A. Asper, L. Zhang, S.-M. Tung, R. Mohan, and L. Dong, "Parotid gland dose in intensity-modulated radiotherapy for head and neck cancer: is what you plan what you get?", International Journal of Radiation Oncology, Biology, Physics **69**, 1290–1296 (2007) (cit. on p. 50).

[64]J. L. Robar, A. Day, J. Clancey, R. Kelly, M. Yewondwossen, H. Hollenhorst, M. Rajaraman, and D. Wilke, "Spatial and dosimetric variability of organs at risk in head-and-neck intensity-modulated radiotherapy", International Journal of Radiation Oncology, Biology, Physics **68**, 1121–1130 (2007) (cit. on p. 50).

[65]C. Han, Y.-J. Chen, A. Liu, T. E. Schultheiss, and J. Y. C. Wong, "Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy", International Journal of Radiation Oncology, Biology, Physics **70**, 1256–1262 (2008) (cit. on p. 50).

[66]P. Castadot, X. Geets, J. A. Lee, N. Christian, and V. Grégoire, "Assessment by a deformable registration method of the volumetric and positional changes of target volumes and organs at risk in pharyngo-laryngeal tumors treated with concomitant chemo-radiation", Radiotherapy and Oncology **95**, 209–217 (2010) (cit. on p. 50).

[67]H. Loo, J. Fairfoul, A. Chakrabarti, J. C. Dean, R. J. Benson, S. J. Jefferies, and N. G. Burnet, "Tumour shrinkage and contour change during radiotherapy increase the dose to organs at risk but not the target volumes for head and neck cancer patients treated on the TomoTherapy HiArt™ system", Clinical Oncology **23**, 40–47 (2011) (cit. on p. 50).

[68]S. Marzi, P. Pinnarò, D. D'Alessio, L. Strigari, V. Bruzzaniti, C. Giordano, G. Giovinazzo, and L. Marucci, "Anatomical and dose changes of gross tumour volume and parotid glands for head and neck cancer patients during intensity-modulated radiotherapy: effect on the probability of xerostomia incidence", Clinical Oncology **24**, e54–62 (2012) (cit. on p. 50).

[69]D. L. Schwartz, A. S. Garden, J. Thomas, Y. Chen, Y. Zhang, J. Lewin, M. S. Chambers, and L. Dong, "Adaptive radiotherapy for head-and-neck cancer: initial clinical outcomes from a prospective trial", International Journal of Radiation Oncology, Biology, Physics **83**, 986–993 (2012) (cit. on pp. 50, 68).

[70]L. Capelle, M. Mackenzie, C. Field, M. Parliament, S. Ghosh, and R. Scrimger, "Adaptive radiotherapy using helical tomotherapy for head and neck cancer in definitive and postoperative settings: initial results", Clinical Oncology **24**, 208–215 (2012) (cit. on p. 50).

[71]D. L. Schwartz, A. S. Garden, S. J. Shah, G. Chronowski, S. Sejpal, D. I. Rosenthal, Y. Chen, Y. Zhang, L. Zhang, P.-F. Wong, J. A. Garcia, K. Kian Ang, and L. Dong, "Adaptive radiotherapy for head and neck cancer: dosimetric results from a prospective clinical trial", Radiotherapy and Oncology **106**, 80–84 (2013) (cit. on p. 50).

[72]W. Wang, H. Yang, W. Hu, G. Shan, W. Ding, C. Yu, B. Wang, X. Wang, and Q. Xu, "Clinical study of the necessity of replanning before the 25th fraction during the course of intensity-modulated radiotherapy for patients with nasopharyngeal carcinoma",

International Journal of Radiation Oncology, Biology, Physics **77**, 617–621 (2010) (cit. on p. 50).

[73]C. B. Simone, 2nd, D. Ly, T. D. Dan, J. Ondos, H. Ning, A. Belard, J. O'Connell, R. W. Miller, and N. L. Simone, "Comparison of intensity-modulated radiotherapy, adaptive radiotherapy, proton radiotherapy, and adaptive proton radiotherapy for treatment of locally advanced head and neck cancer", Radiotherapy and Oncology **101**, 376–382 (2011) (cit. on pp. 50, 105, 118).

[74]N. Howlader, A. Noone, M. Krapcho, J. Garshell, D. Miller, S. Altekruse, C. Kosary, M. Yu, J. Ruhl, Z. Tatalovich, A. Mariotto, D. Lewis, H. Chen, E. Feuer, and K. Cronin, "Estimated new cancer cases and death for 2015.", SEER Cancer Statistics Review, National Cancer Institute. (2015) (cit. on p. 51).

[75]J. S. A. Belderbos, W. D. Heemsbergen, K. De Jaeger, P. Baas, and J. V. Lebesque, "Final results of a phase I/II dose escalation trial in non-small-cell lung cancer using three-dimensional conformal radiotherapy", International Journal of Radiation Oncology, Biology, Physics **66**, 126–134 (2006) (cit. on p. 51).

[76]M. Machtay, K. Bae, B. Movsas, R. Paulus, E. M. Gore, R. Komaki, K. Albain, W. T. Sause, and W. J. Curran, "Higher biologically effective dose of radiotherapy is associated with improved outcomes for locally advanced non-small cell lung carcinoma treated with chemoradiation: an analysis of the radiation therapy oncology group", International Journal of Radiation Oncology, Biology, Physics **82**, 425–434 (2012) (cit. on p. 51).

[77]K. E. Rosenzweig, J. L. Fox, E. Yorke, H. Amols, A. Jackson, V. Rusch, M. G. Kris, C. C. Ling, and S. A. Leibel, "Results of a phase I dose-escalation study using three-dimensional conformal radiotherapy in the treatment of inoperable nonsmall cell lung carcinoma", Cancer **103**, 2118–2127 (2005) (cit. on p. 51).

[78]J. D. Bradley, J. Moughan, M. V. Graham, R. Byhardt, R. Govindan, J. Fowler, J. A. Purdy, J. M. Michalski, E. Gore, and H. Choy, "A phase I/II radiation dose escalation study with concurrent chemotherapy for patients with inoperable stages I to III non-small-cell lung cancer: phase I results of RTOG 0117", International Journal of Radiation Oncology, Biology, Physics **77**, 367–372 (2010) (cit. on p. 51).

[79]J. Bradley, M. V. Graham, K. Winter, J. A. Purdy, R. Komaki, W. H. Roa, J. K. Ryu, W. Bosch, and B. Emami, "Toxicity and outcome results of RTOG 9311: a phase I-II dose-escalation study using three-dimensional conformal radiotherapy in patients with inoperable non-small-cell lung carcinoma", International Journal of Radiation Oncology, Biology, Physics **61**, 318–328 (2005) (cit. on p. 51).

[80]W. P. Levin, H. Kooy, J. S. Loeffler, and T. F. DeLaney, "Proton beam therapy", British Journal of Cancer **93**, 849–854 (2005) (cit. on p. 51).

[81]Y. Kang, X. Zhang, J. Y. Chang, H. Wang, X. Wei, Z. Liao, R. Komaki, J. D. Cox, P. A. Balter, H. Liu, X. R. Zhu, R. Mohan, and L. Dong, "4D proton treatment planning strategy for mobile lung tumors", International Journal of Radiation Oncology, Biology, Physics **67**, 906–914 (2007) (cit. on p. 51).

[82]C. Bert and M. Durante, "Motion in radiotherapy: particle therapy", Physics in Medicine and Biology **56**, R113–144 (2011) (cit. on p. 51).

[83]M. Engelsman, E. Rietzel, and H. M. Kooy, "Four-dimensional proton treatment planning for lung tumors", International Journal of Radiation Oncology, Biology, Physics **64**, 1589–1595 (2006) (cit. on p. 51).

[84]M. Guckenberger, A. Richter, J. Wilbert, M. Flentje, and M. Partridge, "Adaptive radiotherapy for locally advanced non-small-cell lung cancer does not underdose the microscopic disease and has the potential to increase tumor control", International Journal of Radiation Oncology, Biology, Physics **81**, e275–282 (2011) (cit. on p. 51).

[85]Y. Qin, F. Zhang, D. S. Yoo, C. R. Kelsey, F.-F. Yin, and J. Cai, "Adaptive stereotactic body radiation therapy planning for lung cancer", International Journal of Radiation Oncology, Biology, Physics **87**, 209–215 (2013) (cit. on pp. 51, 118).

[86]D. R. Gomez, J. Y. Chang, D. R. Gomez, and J. Y. Chang, "Adaptive radiation for lung cancer, adaptive radiation for lung cancer", Journal of Oncology **2011, 2011**, e898391 (2010) (cit. on pp. 51, 118).

[87]P. Castadot, J. A. Lee, A. Parraga, X. Geets, B. Macq, and V. Grégoire, "Comparison of 12 deformable registration strategies in adaptive radiation therapy for the treatment of head and neck tumors", Radiotherapy and Oncology **89**, 1–12 (2008) (cit. on pp. 55, 80, 166).

[88]D. Yang, S. R. Chaudhari, S. M. Goddu, D. Pratt, D. Khullar, J. O. Deasy, and I. El Naqa, "Deformable registration of abdominal kilovoltage treatment planning CT and tomotherapy daily megavoltage CT for treatment adaptation", Medical Physics **36**, 329–338 (2009) (cit. on pp. 55, 68).

[89]P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion", IEEE Transactions on Pattern Analysis and Machine Intelligence **12**, 629–639 (1990) (cit. on p. 57).

[90]G. Gerig, O. Kubler, R. Kikinis, and F. Jolesz, "Nonlinear anisotropic filtering of MRI data", IEEE Transactions on Medical Imaging **11**, 221–232 (1992) (cit. on p. 57).

[91]V. Grégoire, P. Levendag, K. K. Ang, J. Bernier, M. Braaksma, V. Budach, C. Chao, E. Coche, J. S. Cooper, G. Cosnard, A. Eisbruch, S. El-Sayed, B. Emami, C. Grau, M. Hamoir, et al., "CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC,RTOG consensus guidelines", Radiotherapy and Oncology **69**, 227–236 (2003) (cit. on p. 60).

[92]J. Van Dyk, R. B. Barnett, J. E. Cygler, and P. C. Shragge, "Commissioning and quality assurance of treatment planning computers", International Journal of Radiation Oncology, Biology, Physics **26**, 261–273 (1993) (cit. on p. 61).

[93]D. A. Low, "Gamma dose distribution evaluation tool", Journal of Physics: Conference Series **250**, 012071 (2010) (cit. on p. 61).

[94]D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions", Medical Physics **25**, 656–661 (1998) (cit. on p. 61).

[95]Y. Mishchenko, *3D euclidean distance transform for variable data aspect ratio*, 2011 (cit. on p. 62).

[96]J. Scherman, "Development and evaluation of methods for comparison of dose distributions in radiotherapy using calculated, synthetic and simulated measured dose distributions.", PhD thesis (Lund University, 2009) (cit. on p. 62).

[97]A. Richter, Q. Hu, D. Steglich, K. Baier, J. Wilbert, M. Guckenberger, and M. Flentje, "Investigation of the usability of conebeam CT data sets for dose calculation", Radiation Oncology **3**, 42 (2008) (cit. on pp. 68, 71, 150).

[98]K. Usui, Y. Ichimaru, Y. Okumura, K. Murakami, M. Seo, E. Kunieda, and K. Ogawa, "Dose calculation with a cone beam CT image in image-guided radiation therapy", Radiological Physics and Technology **6**, 107–114 (2013) (cit. on pp. 68, 71).

[99]H. Guan and H. Dong, "Dose calculation accuracy using cone-beam CT (CBCT) for pelvic adaptive radiotherapy", Physics in Medicine and Biology **54**, 6239–6250 (2009) (cit. on pp. 68, 81, 150).

[100]H. Zhong, J. Kim, and I. J. Chetty, "Analysis of deformable image registration accuracy using computational modeling", Medical Physics **37**, 970–979 (2010) (cit. on pp. 68, 99).

[101]J. D. Lawson, E. Schreibmann, A. B. Jani, and T. Fox, "Quantitative evaluation of a cone beam computed tomography (CBCT)-CT deformable image registration method for adaptive radiation therapy", Journal of Applied Clinical Medical Physics **8** (2007) (cit. on p. 68).

[102]J. Hou, M. Guerrero, W. Chen, and W. D. D'Souza, "Deformable planning CT to cone-beam CT image registration in head-and-neck cancer", Medical Physics **38**, 2088–2094 (2011) (cit. on p. 68).

[103]M. Modat, J. McClelland, and S. Ourselin, "Lung registration using the NiftyReg package", Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (Beijing, China, 2010) (2010) (cit. on pp. 69, 87).

[104]G. X. Ding, D. M. Duggan, C. W. Coffey, M. Deeley, D. E. Hallahan, A. Cmelak, and A. Malcolm, "A study on adaptive IMRT treatment planning using kV cone-beam CT", Radiotherapy and Oncology **85**, 116–125 (2007) (cit. on p. 70).

[105]J. Annkah, N. Hindocha, S. Moinuddin, K. Ricketts, I. Rosenberg, G. Royle, and A. Adeyemi, "Assessment of the dosimetric accuracies of CATPhan 504 and CIRS 062 using kV-CBCT for performing direct calculations", Journal of Medical Physics **39**, 133 (2014) (cit. on p. 73).

[106]The Phantom Laboratory, *Catphan 504 manual*, 2013 (cit. on p. 73).

[107]D. A. Jaffray and J. H. Siewerdsen, "Cone-beam computed tomography with a flat-panel imager: initial performance characterization", Medical Physics **27**, 1311–1323 (2000) (cit. on p. 81).

[108]H. Kato, M. Kanematsu, O. Tanaka, K. Mizuta, M. Aoki, T. Shibata, T. Yamashita, Y. Hirose, and H. Hoshi, "Head and neck squamous cell carcinoma: usefulness of diffusion-weighted MR imaging in the prediction of a neoadjuvant therapeutic effect", European Radiology **19**, 103–109 (2009) (cit. on p. 82).

[109]H. C. Thoeny, F. De Keyzer, and A. D. King, "Diffusion-weighted MR imaging in the head and neck", Radiology **263**, 19–32 (2012) (cit. on p. 82).

[110]J. Kim, M. M. Matuszak, K. Saitou, and J. M. Balter, "Distance-preserving rigidity penalty on deformable image registration of multiple skeletal components in the neck", Medical Physics **40**, 121907 (2013) (cit. on pp. 82, 100, 115, 149).

[111]P. Cachier, E. Bardinet, D. Dormont, X. Pennec, and N. Ayache, "Iconic feature based nonrigid registration: the PASHA algorithm", Computer Vision and Image Understanding **89**, 272–298 (2003) (cit. on pp. 82, 162, 163).

[112]R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A. K. Garg, and T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets", Physics in Medicine and Biology **54**, 1849–1870 (2009) (cit. on p. 86).

[113]P. Risholm, J. Balter, and W. M. Wells, "Estimation of delivered dose in radiotherapy: the influence of registration uncertainty", Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (Toronto, Canada, 2011) **14**, 548–555 (2011) (cit. on p. 86).

[114]E. T. Bender and W. A. Tomé, "The utilization of consistency metrics for error analysis in deformable image registration", Physics in Medicine and Biology **54**, 5561–5577 (2009) (cit. on p. 86).

[115]R. Varadhan, G. Karangelis, K. Krishnan, and S. Hui, "A framework for deformable image registration validation in radiotherapy clinical applications", Journal of Applied Clinical Medical Physics **14**, 4066 (2013) (cit. on pp. 86, 90).

[116]M. Hub, M. L. Kessler, and C. P. Karger, "A stochastic approach to estimate the uncertainty involved in b-spline image registration", IEEE Transactions on Medical Imaging **28**, 1708–1716 (2009) (cit. on p. 86).

[117]A. Cherpak, M. Serban, J. Seuntjens, and J. E. Cygler, "4D dose-position verification in radiation therapy using the RADPOS system in a deformable lung phantom", Medical Physics **38**, 179–187 (2011) (cit. on p. 86).

[118]U. J. Yeo, M. L. Taylor, J. R. Supple, R. L. Smith, L. Dunn, T. Kron, and R. D. Franich, "Is it sensible to "deform" dose? 3D experimental validation of dose-warping", Medical Physics **39**, 5065–5072 (2012) (cit. on p. 86).

[119]J. Pukala, S. L. Meeks, R. J. Staton, F. J. Bova, R. R. Mañon, and K. M. Langen, "A virtual phantom library for the quantification of deformable image registration uncertainties in patients with cancers of the head and neck", Medical Physics **40**, 111703 (2013) (cit. on p. 86).

[120]N. K. Saleh-Sayah, E. Weiss, F. J. Salguero, and J. V. Siebers, "A distance to dose difference tool for estimating the required spatial accuracy of a displacement vector field", Medical Physics **38**, 2318–2323 (2011) (cit. on pp. 86, 91, 96, 98).

[121]D. Tilly, N. Tilly, and A. Ahnesjö, "Dose mapping sensitivity to deformable registration uncertainties in fractionated radiotherapy – applied to prostate proton treatments", BMC Medical Physics **13**, 1–12 (2013) (cit. on p. 86).

[122]M. Rosu, I. J. Chetty, J. M. Balter, M. L. Kessler, D. L. McShan, and R. K. Ten Haken, "Dose reconstruction in deforming lung anatomy: dose grid size effects and clinical implications", Medical Physics **32**, 2487–2495 (2005) (cit. on p. 86).

[123]F. J. Salguero, N. K. Saleh-Sayah, C. Yan, and J. V. Siebers, "Estimation of three-dimensional intrinsic dosimetric uncertainties resulting from using deformable image registration for dose mapping", Medical Physics **38**, 343–353 (2011) (cit. on pp. 86, 97).

[124]M. J. Murphy, F. J. Salguero, J. V. Siebers, D. Staub, and C. Vaman, "A method to estimate the effect of deformable image registration uncertainties on daily dose mapping", Medical Physics **39**, 573–580 (2012) (cit. on pp. 86, 96).

[125]H. Paganetti, H. Jiang, J. A. Adams, G. T. Chen, and E. Rietzel, "Monte Carlo simulations with time-dependent geometries to investigate effects of organ motion with high temporal resolution", International Journal of Radiation Oncology, Biology, Physics **60**, 942–950 (2004) (cit. on p. 86).

[126]E. Heath and J. Seuntjens, "A direct voxel tracking method for four-dimensional Monte Carlo dose calculations in deforming anatomy", Medical Physics **33**, 434–445 (2006) (cit. on p. 86).

[127]H. Zhong and J. V. Siebers, "Monte Carlo dose mapping on deforming anatomy", Physics in Medicine and Biology **54**, 5815–5830 (2009) (cit. on p. 86).

[128]B. Schaly, J. A. Kempe, G. S. Bauman, J. J. Battista, and J. Van Dyk, "Tracking the dose distribution in radiation therapy by accounting for variable anatomy", Physics in Medicine and Biology **49**, 791–805 (2004) (cit. on p. 86).

[129]A. Lourenço, "Deformable registration, dose remapping and summation for head and neck adaptive radiotherapy applications", MSc Thesis (University of Lisbon, 2013) (cit. on p. 86).

[130]T. E. Schultheiss, W. A. Tomé, and C. G. Orton, "It is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy", Medical Physics **39**, 6531–6533 (2012) (cit. on p. 99).

[131]U. J. Yeo, J. R. Supple, M. L. Taylor, R. Smith, T. Kron, and R. D. Franich, "Performance of 12 DIR algorithms in low-contrast regions for mass and density conserving deformation", Medical Physics **40**, 101701 (2013) (cit. on p. 99).

[132]T. E. Schultheiss and W. A. Tome, "Response to "comment on 'it is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy'" [med. phys.39, 6531–6533 (2012)]", Medical Physics **40**, 017102 (2013) (cit. on p. 99).

[133]S. Nithiananthan, S. Schafer, D. J. Mirota, J. W. Stayman, W. Zbijewski, D. D. Reh, G. L. Gallia, and J. H. Siewerdsen, "Extra-dimensional demons: a method for incorporating missing tissue in deformable image registration", Medical Physics **39**, 5718–5731 (2012) (cit. on p. 100).

[134]J. E. Turner, *Atoms, radiation, and radiation protection*, 3rd edition (Wiley-VCH, Weinheim, 2007), 606 pp. (cit. on p. 102).

[135]International Commission on Radiation Units and Measurements, *Clinical proton dosimetry part I: beam production, beam delivery and measurement of absorbed dose (ICRU*

report 59) (International Commission on Radiation Units and Measurements (ICRU), 1998), p. 3123 (cit. on p. 102).

[136]A. E. Nahum, D. P. Dearnaley, and G. G. Steel, "Prospects for proton-beam radiotherapy", European Journal of Cancer (1990) **30A**, 1577–1583 (1994) (cit. on p. 103).

[137]B. M. Beadle, K.-P. Liao, L. S. Elting, T. A. Buchholz, K. K. Ang, A. S. Garden, and B. A. Guadagnolo, "Improved survival using intensity-modulated radiation therapy in head and neck cancers: a SEER-medicare analysis", Cancer **120**, 702–710 (2014) (cit. on p. 103).

[138]S. Kandula, X. Zhu, A. S. Garden, M. Gillin, D. I. Rosenthal, K.-K. Ang, R. Mohan, M. V. Amin, J. A. Garcia, R. Wu, N. Sahoo, and S. J. Frank, "Spot-scanning beam proton therapy vs intensity-modulated radiation therapy for ipsilateral head and neck malignancies: a treatment planning comparison", Medical Dosimetry **38**, 390–394 (2013) (cit. on pp. 103, 106).

[139]A. J. Lomax, M. Goitein, and J. Adams, "Intensity modulation in radiotherapy: photons versus protons in the paranasal sinus", Radiotherapy and Oncology **66**, 11–18 (2003) (cit. on p. 103).

[140]M. Steneker, A. Lomax, and U. Schneider, "Intensity modulated photon and proton therapy for the treatment of head and neck tumors", Radiotherapy and Oncology **80**, 263–267 (2006) (cit. on p. 103).

[141]E. M. Quan, W. Liu, R. Wu, Y. Li, S. J. Frank, X. Zhang, X. R. Zhu, and R. Mohan, "Preliminary evaluation of multifield and single-field optimization for the treatment planning of spot-scanning proton therapy of head and neck cancer", Medical Physics **40**, 081709 (2013) (cit. on pp. 103, 107).

[142]A. W. Chan, L. J. Liebsch, D. G. Deschler, J. A. Adams, L. V. Vrishali, J. F. McIntyre, P. Pommier, R. L. Fabian, and P. M. Busse, "Proton radiotherapy for T4 nasopharyngeal carcinoma", ASCO Meeting Abstracts **22**, 5574 (2004) (cit. on p. 103).

[143]J. D. Slater, L. T. Yonemoto, D. W. Mantik, D. A. Bush, W. Preston, R. I. Grove, D. W. Miller, and J. M. Slater, "Proton radiation for treatment of cancer of the oropharynx: early experience at Loma Linda University Medical Center using a concomitant boost technique", International Journal of Radiation Oncology, Biology, Physics **62**, 494–500 (2005) (cit. on p. 103).

[144]H. Nishimura, T. Ogino, M. Kawashima, K. Nihei, S. Arahira, M. Onozawa, S. Katsuta, and T. Nishio, "Proton-beam therapy for olfactory neuroblastoma", International Journal of Radiation Oncology, Biology, Physics **68**, 758–762 (2007) (cit. on pp. 103, 104).

[145]V. A. Resto, A. W. Chan, D. G. Deschler, and D. T. Lin, "Extent of surgery in the management of locally advanced sinonasal malignancies", Head & Neck **30**, 222–229 (2008) (cit. on p. 103).

[146]S. J. Frank, J. D. Cox, M. Gillin, R. Mohan, A. S. Garden, D. I. Rosenthal, G. B. Gunn, R. S. Weber, M. S. Kies, J. S. Lewin, M. F. Munsell, M. B. Palmer, N. Sahoo, X. Zhang, W. Liu, et al., "Multifield optimization intensity modulated proton therapy for head and

neck tumors: a translation to practice", International Journal of Radiation Oncology, Biology, Physics **89**, 846–853 (2014) (cit. on pp. 103, 107).

[147]A. M. Allen, T. Pawlicki, L. Dong, E. Fourkal, M. Buyyounouski, K. Cengel, J. Plastaras, M. K. Bucci, T. I. Yock, L. Bonilla, R. Price, E. E. Harris, and A. A. Konski, "An evidence based review of proton beam therapy: the report of ASTRO's emerging technology committee", Radiotherapy and Oncology **103**, 8–11 (2012) (cit. on p. 104).

[148]W. Schneider, T. Bortfeld, and W. Schlegel, "Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions", Physics in Medicine and Biology **45**, 459–478 (2000) (cit. on p. 105).

[149]U. Amaldi, A. Bianchi, Y.-H. Chang, A. Go, W. Hajdas, N. Malakhov, J. Samarati, F. Sauli, and D. Watts, "Construction, test and operation of a proton range radiography system", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **629**, 337–344 (2011) (cit. on p. 105).

[150]X. Zhu and G. El Fakhri, "Proton therapy verification with PET imaging", Theranostics **3**, 731–740 (2013) (cit. on p. 105).

[151]J. M. Verburg, K. Riley, T. Bortfeld, and J. Seco, "Energy- and time-resolved detection of prompt gamma-rays for proton range verification", Physics in Medicine and Biology **58**, L37–49 (2013) (cit. on p. 105).

[152]S. Shimizu, T. Matsuura, M. Umezawa, K. Hiramoto, N. Miyamoto, K. Umegaki, and H. Shirato, "Preliminary analysis for integration of spot-scanning proton beam therapy and real-time imaging and gating", Physica Medica (2014) (cit. on p. 105).

[153]J. Seabra, S. Brousmiche, R. Labarbe, M. Vila, S. Rit, D. Wikler, J. Lee, K. Teo, J. Orban de Xivry, and B. Macq, "Design of cone-beam CT for proton therapy gantry", (2014) (cit. on p. 105).

[154]B. Winey, Y. Park, M. Zhu, H.-M. Lu, M. Moteabbed, and G. Sharp, "Proton gantry mounted CBCT development at massachusetts general hospital", (2014) (cit. on p. 105).

[155]S. Takao, S. Shimizu, N. Miyamoto, T. Matsuura, C. Toramatsu, H. Nihongi, T. Yamada, K. Matsuda, T. Sasaki, Y. Nagamine, R. Raba, T. Umekawa, K. Umegaki, and H. Shirato, "Commissioning of the on-board cone-beam CT system equipped on the rotating gantry of a proton therapy system", (2014) (cit. on p. 105).

[156]A. Bolsi, A. J. Lomax, E. Pedroni, G. Goitein, and E. Hug, "Experiences at the paul scherrer institute with a remote patient positioning procedure for high-throughput proton radiation therapy", International Journal of Radiation Oncology, Biology, Physics **71**, 1581–1590 (2008) (cit. on p. 105).

[157]N. J. Lomax and S. G. Scheib, "Quantifying the degree of conformity in radiosurgery treatment planning", International Journal of Radiation Oncology, Biology, Physics **55**, 1409–1419 (2003) (cit. on p. 107).

[158]X. Wang, X. Zhang, L. Dong, H. Liu, M. Gillin, A. Ahamad, K. Ang, and R. Mohan, "Effectiveness of noncoplanar IMRT planning using a parallelized multiresolution beam

angle optimization method for paranasal sinus carcinoma", International Journal of Radiation Oncology, Biology, Physics **63**, 594–601 (2005) (cit. on p. 107).

[159]M. D'Arienzo, S. G. Masciullo, V. d. Sanctis, M. F. Osti, L. Chiacchiararelli, and R. M. Enrici, "Integral dose and radiation-induced secondary malignancies: comparison between stereotactic body radiation therapy and three-dimensional conformal radiotherapy", International Journal of Environmental Research and Public Health **9**, 4223–4240 (2012) (cit. on p. 107).

[160]A. Bharatha, M. Hirose, N. Hata, S. K. Warfield, M. Ferrant, K. H. Zou, E. Suarez-Santana, J. Ruiz-Alzola, A. D'Amico, R. A. Cormack, R. Kikinis, F. A. Jolesz, and C. M. Tempany, "Evaluation of three-dimensional finite element-based deformable registration of pre- and intraoperative prostate imaging", Medical Physics **28**, 2551–2560 (2001) (cit. on p. 115).

[161]A. Al-Mayah, J. Moseley, S. Hunter, M. Velec, L. Chau, S. Breen, and K. Brock, "Biomechanical-based image registration for head and neck radiation treatment", Physics in Medicine and Biology **55**, 6491–6500 (2010) (cit. on p. 115).

[162]M. Stuschke, A. Kaiser, J. A. Jawad, C. Pöttgen, S. Levegrün, and J. Farr, "Multi-scenario based robust intensity-modulated proton therapy (IMPT) plans can account for set-up errors more effectively in terms of normal tissue sparing than planning target volume (PTV) based intensity-modulated photon plans in the head and neck region", Radiation Oncology **8**, 145 (2013) (cit. on p. 115).

[163]J. Seabra, S. Brousmich, R. Labarbe, J. Lee, and B. Macq, "Methodological study of geometric deformation for CBCT in proton therapy gantry", Proceedings of the Annual Symposium of the IEEE EMBS Benelux Chapter (Belgium, 2011) (2011) (cit. on p. 116).

[164]H. Paganetti, "Range uncertainties in proton therapy and the role of Monte Carlo simulations", Physics in Medicine and Biology **57**, R99–R117 (2012) (cit. on p. 116).

[165]C. B. Simone and R. Rengan, "The use of proton therapy in the treatment of lung cancers", Cancer Journal **20**, 427–432 (2014) (cit. on p. 118).

[166]E. Roelofs, M. Engelsman, C. Rasch, L. Persoon, S. Qamhiyeh, D. de Ruysscher, F. Verhaegen, M. Pijls-Johannesma, P. Lambin, and ROCOCO Consortium, "Results of a multicentric in silico clinical trial (ROCOCO): comparing radiotherapy with photons and protons for non-small cell lung cancer", Journal of Thoracic Oncology **7**, 165–176 (2012) (cit. on p. 118).

[167]J. Y. Chang, X. Zhang, X. Wang, Y. Kang, B. Riley, S. Bilton, R. Mohan, R. Komaki, and J. D. Cox, "Significant reduction of normal tissue dose by proton radiotherapy compared with three-dimensional conformal or intensity-modulated radiation therapy in stage I or stage III non-small-cell lung cancer", International Journal of Radiation Oncology, Biology, Physics **65**, 1087–1096 (2006) (cit. on p. 118).

[168]A. H. Kesarwala, C. J. Ko, H. Ning, E. Xanthopoulos, K. E. Haglund, W. P. O'Meara, C. B. Simone, and R. Rengan, "Intensity-modulated proton therapy for elective nodal irradiation and involved-field radiation in the definitive treatment of locally advanced

non-small-cell lung cancer: a dosimetric study", Clinical Lung Cancer **16**, 237–244 (2015) (cit. on p. 118).

[169]C. E. DeSantis, C. C. Lin, A. B. Mariotto, R. L. Siegel, K. D. Stein, J. L. Kramer, R. Alteri, A. S. Robbins, and A. Jemal, "Cancer treatment and survivorship statistics, 2014", CA: A Cancer Journal for Clinicians **64**, 252–271 (2014) (cit. on p. 118).

[170]K. C. J. Wink, E. Roelofs, T. Solberg, L. Lin, C. B. Simone, A. Jakobi, C. Richter, P. Lambin, and E. G. C. Troost, "Particle therapy for non-small cell lung tumors: where do we stand? a systematic review of the literature", Frontiers in Oncology **4** (2014) (cit. on p. 118).

[171]M. Guckenberger, J. Wilbert, A. Richter, K. Baier, and M. Flentje, "Potential of adaptive radiotherapy to escalate the radiation dose in combined radiochemotherapy for locally advanced non-small cell lung cancer", International Journal of Radiation Oncology, Biology, Physics **79**, 901–908 (2011) (cit. on p. 118).

[172]M. van Zwienen, S. van Beek, J. Belderbos, S. van Kranen, C. Rasch, M. van Herk, and J. Sonke, "Anatomical changes during radiotherapy of lung cancer patients", International Journal of Radiation Oncology, Biology, Physics **72**, S111 (2008) (cit. on p. 118).

[173]K. R. Britton, G. Starkschall, S. L. Tucker, T. Pan, C. Nelson, J. Y. Chang, J. D. Cox, R. Mohan, and R. Komaki, "Assessment of gross tumor volume regression and motion changes during radiotherapy for non-small-cell lung cancer as measured by four-dimensional computed tomography", International Journal of Radiation Oncology, Biology, Physics **68**, 1036–1046 (2007) (cit. on p. 118).

[174]C. Ozhasoglu and M. J. Murphy, "Issues in respiratory motion compensation during external-beam radiotherapy", International Journal of Radiation Oncology, Biology, Physics **52**, 1389–1399 (2002) (cit. on p. 118).

[175]A. T. Berman, S. S. James, and R. Rengan, "Proton beam therapy for non-small cell lung cancer: current clinical evidence and future directions", Cancers **7**, 1178–1190 (2015) (cit. on p. 118).

[176]Z. Hui, X. Zhang, G. Starkschall, Y. Li, R. Mohan, R. Komaki, J. D. Cox, and J. Y. Chang, "Effects of interfractional motion and anatomic changes on proton therapy dose distribution in lung cancer", International Journal of Radiation Oncology, Biology, Physics **72**, 1385–1395 (2008) (cit. on p. 118).

[177]S. Rit, M. V. Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. C. Sharp, "The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)", Journal of Physics: Conference Series **489**, 012079 (2014) (cit. on pp. 122, 128).

[178]H. Knutsson and M. Andersson, "Morphons: segmentation using elastic canvas and paint on priors", **2**, II–1226–9 (2005) (cit. on p. 123).

[179]J.-J. Sonke and J. Belderbos, "Adaptive radiotherapy for lung cancer", Seminars in Radiation Oncology **20**, 94–106 (2010) (cit. on p. 124).

[180]G. Bernard, M. Verleysen, and J. A. Lee, "Incremental classification of objects in scenes: application to the delineation of images", Neurocomputing **152**, 45–57 (2015) (cit. on p. 124).

[181]J. Cousty, G. Bertrand, L. Najman, and M. Couprie, "Watershed cuts: minimum spanning forests and the drop of water principle", IEEE Transactions on Pattern Analysis and Machine Intelligence **31**, 1362–1374 (2009) (cit. on p. 124).

[182]Y. Freund and R. E. Schapire, "Large margin classification using the perceptron algorithm", Mach. Learn. **37**, 277–296 (1999) (cit. on p. 124).

[183]S. I. Gallant, "Perceptron-based learning algorithms", IEEE Transactions on Neural Networks **1**, 179–191 (1990) (cit. on p. 124).

[184]U. Schneider, E. Pedroni, and A. Lomax, "The calibration of CT hounsfield units for radiotherapy treatment planning", Physics in Medicine and Biology **41**, 111–124 (1996) (cit. on p. 125).

[185]P. C. Park, J. Cheung, X. R. Zhu, N. Sahoo, L. Court, and L. Dong, "Fast range-corrected proton dose approximation method using prior dose distribution", Physics in Medicine and Biology **57**, 3555–3569 (2012) (cit. on p. 125).

[186]M. Urie, M. Goitein, and M. Wagner, "Compensating for heterogeneities in proton radiation therapy", Physics in Medicine and Biology **29**, 553–566 (1984) (cit. on p. 126).

[187]B. Ohnesorge, T. Flohr, and K. Klingenbeck-Regn, "Efficient object scatter correction algorithm for third and fourth generation CT scanners", European Radiology **9**, 563–569 (1999) (cit. on p. 128).

[188]H. Li, R. Mohan, and X. R. Zhu, "Scatter kernel estimation with an edge-spread function method for cone-beam computed tomography imaging", Physics in Medicine and Biology **53**, 6729–6748 (2008) (cit. on p. 128).

[189]M. Sun and J. M. Star-Lack, "Improved scatter correction using adaptive scatter kernel superposition", Physics in Medicine and Biology **55**, 6695–6720 (2010) (cit. on p. 128).

[190]S. Jan, G. Santin, D. Strul, S. Staelens, K. Assié, D. Autret, S. Avner, R. Barbier, M. Bardiès, P. M. Bloomfield, D Brasse, V. Breton, P. Bruyndonckx, I. Buvat, A. F. Chatziioannou, et al., "GATE: a simulation toolkit for PET and SPECT", Physics in Medicine and Biology **49**, 4543 (2004) (cit. on p. 129).

[191]D. A. Palma, J. van Sörnsen de Koste, W. F. A. R. Verbakel, A. Vincent, and S. Senan, "Lung density changes after stereotactic radiotherapy: a quantitative analysis in 50 patients", International Journal of Radiation Oncology, Biology, Physics **81**, 974–978 (2011) (cit. on p. 149).

[192]L. Risser, M. Heinrich, T. Matin, and J. Schnabel, "Piecewise-diffeomorphic registration of 3D CT/MR pulmonary images with sliding conditions", 1351–1354 (2012) (cit. on p. 149).

[193]C. Kurz, G. Dedes, A. Resch, M. Reiner, U. Ganswindt, R. Nijhuis, C. Thieke, C. Belka, K. Parodi, and G. Landry, "Comparing cone-beam CT intensity correction methods for dose recalculation in adaptive intensity-modulated photon and proton therapy for head and neck cancer", Acta Oncologica **54**, 1651–1657 (2015) (cit. on p. 150).

[194]J. Li, W. Yao, Y. Xiao, and Y. Yu, "Feasibility of improving cone-beam CT number consistency using a scatter correction algorithm", Journal of Applied Clinical Medical Physics **14** (2013) (cit. on p. 150).

[195]Y.-K. Park, G. C. Sharp, J. Phillips, and B. A. Winey, "Proton dose calculation on scatter-corrected CBCT image: feasibility study for adaptive proton therapy", Medical Physics **42**, 4449–4459 (2015) (cit. on p. 150).

[196]X. Jia, J. Schümann, H. Paganetti, and S. B. Jiang, "GPU-based fast Monte Carlo dose calculation for proton therapy", Physics in Medicine and Biology **57**, 7783 (2012) (cit. on p. 150).

[197]H. Nyström, "The role of protons in modern and biologically-guided radiotherapy", Acta Oncologica **49**, 1124–1131 (2010) (cit. on p. 156).

[198]T. E. Yankeelov and J. C. Gore, "Preface to the special issue on quantitative imaging in cancer", Magnetic Resonance Imaging **30**, 1201–1202 (2012) (cit. on pp. 156, 157).

[199]C. D. Marcus, V. Ladam-Marcus, C. Cucu, O. Bouché, L. Lucas, and C. Hoeffel, "Imaging techniques to evaluate the response to treatment in oncology: current standards and perspectives", Critical Reviews in Oncology/Hematology **72**, 217–238 (2009) (cit. on pp. 156, 167).

[200]P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. P. M. van Stiphout, P. Granton, C. M. L. Zegers, R. Gillies, R. Boellard, A. Dekker, and H. J. W. L. Aerts, "Radiomics: extracting more information from medical images using advanced feature analysis", European Journal of Cancer **48**, 441–446 (2012) (cit. on p. 156).

[201]P. Dirix, V. Vandecaveye, F. De Keyzer, S. Stroobants, R. Hermans, and S. Nuyts, "Dose painting in radiotherapy for head and neck squamous cell carcinoma: value of repeated functional imaging with (18)f-FDG PET, (18)f-fluoromisonidazole PET, diffusion-weighted MRI, and dynamic contrast-enhanced MRI", Journal of Nuclear Medicine **50**, 1020–1027 (2009) (cit. on pp. 156, 167).

[202]A. R. Padhani and K. A. Miles, "Multiparametric imaging of tumor response to therapy", Radiology **256**, 348–364 (2010) (cit. on p. 157).

[203]D. C. Colvin, M. E. Loveless, M. D. Does, Z. Yue, T. E. Yankeelov, and J. C. Gore, "Earlier detection of tumor treatment response using magnetic resonance diffusion imaging with oscillating gradients", Magnetic Resonance Imaging **29**, 315–323 (2011) (cit. on p. 157).

[204]J. J. W. Lagendijk, B. W. Raaymakers, A. J. E. Raaijmakers, J. Overweg, K. J. Brown, E. M. Kerkhof, R. W. van der Put, B. Hårdemark, M. van Vulpen, and U. A. van der Heide, "MRI/linac integration", Radiotherapy and Oncology **86**, 25–29 (2008) (cit. on p. 157).

[205]B. G. Fallone, B. Murray, S. Rathee, T. Stanescu, S. Steciw, S. Vidakovic, E. Blosser, and D. Tymofichuk, "First MR images obtained during megavoltage photon irradiation from a prototype integrated linac-MR system", Medical Physics **36**, 2084–2088 (2009) (cit. on p. 157).

[206]W. H. Organization, "WHO handbook for reporting results of cancer treatment", (1979) (cit. on p. 157).

[207]A. B. Miller, B. Hoogstraten, M. Staquet, and A. Winkler, "Reporting results of cancer treatment", Cancer **47**, 207–214 (1981) (cit. on p. 157).

[208]B. Zhao, L. H. Schwartz, and S. M. Larson, "Imaging surrogates of tumor response to therapy: anatomic and functional biomarkers", Journal of Nuclear Medicine **50**, 239–249 (2009) (cit. on p. 157).

[209]P. Therasse, S. G. Arbuck, E. A. Eisenhauer, J. Wanders, R. S. Kaplan, L. Rubinstein, J. Verweij, M. Van Glabbeke, A. T. van Oosterom, M. C. Christian, and S. G. Gwyther, "New guidelines to evaluate the response to treatment in solid tumors. european organization for research and treatment of cancer, national cancer institute of the united states, national cancer institute of canada", Journal of the National Cancer Institute **92**, 205–216 (2000) (cit. on p. 157).

[210]V. Vandecaveye, F. De Keyzer, S. Nuyts, K. Deraedt, P. Dirix, P. Hamaekers, V. Vander Poorten, P. Delaere, and R. Hermans, "Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings", International Journal of Radiation Oncology, Biology, Physics **67**, 960–971 (2007) (cit. on pp. 157, 170).

[211]S. K. Das and R. K. Ten Haken, "Functional and molecular image guidance in radiotherapy treatment planning optimization", Seminars in radiation oncology **21**, 111–118 (2011) (cit. on p. 158).

[212]U. A. van der Heide, A. C. Houweling, G. Groenendaal, R. G. H. Beets-Tan, and P. Lambin, "Functional MRI for radiotherapy dose painting", Magnetic Resonance Imaging **30**, 1216–1223 (2012) (cit. on p. 158).

[213]S. Leibfarth, D. Mönnich, S. Welz, C. Siegel, N. Schwenzer, H. Schmidt, D. Zips, and D. Thorwarth, "A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning", Acta Oncologica **52**, 1353–1359 (2013) (cit. on pp. 158, 166).

[214]P. Slagmolen, D. Loeckx, S. Roels, X. Geets, F. Maes, K. Haustermans, and P. Suetens, "Nonrigid registration of multitemporal CT and MR images for radiotherapy treatment planning", Biomedical Image Registration, 297–305 (2006) (cit. on pp. 158, 166).

[215]A. du Bois d'Aische, M. De Craene, X. Geets, V. Grégoire, B. Macq, and S. K. Warfield, "Estimation of the deformations induced by articulated bodies: registration of the spinal column", Biomedical Signal Processing and Control **2**, 16–24 (2007) (cit. on p. 158).

[216]M. Söhn, M. Birkner, Y. Chi, J. Wang, Y. Di, B. Berger, and M. Alber, "Model-independent, multimodality deformable image registration by local matching of anatomical features and minimization of elastic energy", Medical Physics **35**, 866–878 (2008) (cit. on p. 158).

[217]K. K. Brock and Deformable Registration Accuracy Consortium, "Results of a multi-institution deformable registration accuracy study (MIDRAS)", International Journal of Radiation Oncology, Biology, Physics **76**, 583–596 (2010) (cit. on p. 158).

[218]A. Akbarzadeh, D. Gutierrez, A. Baskin, M. R. Ay, A. Ahmadian, N. Riahi Alam, K. O. Lövblad, and H. Zaidi, "Evaluation of whole-body MR to CT deformable image registration", Journal of Applied Clinical Medical Physics **14**, 4163 (2013) (cit. on p. 158).

[219]I. D. Dmitriev, C. E. Loo, W. V. Vogel, K. E. Pengel, and K. G. A. Gilhuijs, "Fully automated deformable registration of breast DCE-MRI and PET/CT", Physics in Medicine and Biology **58**, 1221–1233 (2013) (cit. on p. 158).

[220]N. Tustison, B. Avants, P. Cook, Y. Zheng, A. Egan, P. Yushkevich, and J. Gee, "N4ITK: improved N3 bias correction", IEEE Transactions on Medical Imaging **29**, 1310–1320 (2010) (cit. on p. 163).

[221]C. Veiga, J. McClelland, S. Moinuddin, A. Lourenço, K. Ricketts, J. Annkah, M. Modat, S. Ourselin, D. D'Souza, and G. Royle, "Toward adaptive radiotherapy for head and neck patients: feasibility study on using CT-to-CBCT deformable registration for "dose of the day" calculations", Medical Physics **41**, 031703 (2014) (cit. on p. 166).

[222]A. R. Padhani and A. A. Khan, "Diffusion-weighted (DW) and dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) for monitoring anticancer therapy", Targeted Oncology **5**, 39–52 (2010) (cit. on p. 167).

[223]C. J. Galbán, S. K. Mukherji, T. L. Chenevert, C. R. Meyer, D. A. Hamstra, P. H. Bland, T. D. Johnson, B. A. Moffat, A. Rehemtulla, A. Eisbruch, and B. D. Ross, "A feasibility study of parametric response map analysis of diffusion-weighted magnetic resonance imaging scans of head and neck cancer patients for providing early detection of therapeutic efficacy", Translational oncology **2**, 184–190 (2009) (cit. on pp. 167, 170).

[224]G. K. Rohde, A. S. Barnett, P. J. Basser, S. Marenco, and C. Pierpaoli, "Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI", Magnetic Resonance in Medicine **51**, 103–114 (2004) (cit. on p. 167).

[225]J. R. Teruel, H. E. Fjøsne, A. Ostlie, D. Holland, A. M. Dale, T. F. Bathen, and P. E. Goa, "Inhomogeneous static magnetic field-induced distortion correction applied to diffusion weighted MRI of the breast at 3T", Magnetic Resonance in Medicine (2014) (cit. on p. 167).

[226]B. W. Raaymakers, J. J. W. Lagendijk, J. Overweg, J. G. M. Kok, A. J. E. Raaijmakers, E. M. Kerkhof, R. W. van der Put, I. Meijsing, S. P. M. Crijns, F. Benedosso, M. van Vulpen, C. H. W. de Graaff, J. Allen, and K. J. Brown, "Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept", Physics in Medicine and Biology **54**, N229–237 (2009) (cit. on p. 168).

[227]N. Burgos, M. J. Cardoso, F. Guerreiro, C. Veiga, M. Modat, J. McClelland, A.-C. Knopf, S. Punwani, D. Atkinson, S. R. Arridge, B. F. Hutton, and S. Ourselin, "Robust CT synthesis for radiotherapy planning: application to the head and neck region", Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (Munich, Germany, 2015), 476–484 (2015) (cit. on p. 168).

[228]S. Roy, W.-T. Wang, A. Carass, J. L. Prince, J. A. Butman, and D. L. Pham, "PET attenuation correction using synthetic CT from ultrashort echo-time MRI", Journal of Nuclear Medicine **55**, 2071–2077 (2014) (cit. on p. 168).

[229]T. L. Chenevert, L. D. Stegman, J. M. G. Taylor, P. L. Robertson, H. S. Greenberg, A. Rehemtulla, and B. D. Ross, "Diffusion magnetic resonance imaging: an early surrogate marker of therapeutic efficacy in brain tumors", Journal of the National Cancer Institute **92**, 2029–2036 (2000) (cit. on p. 170).

[230]B. A. Moffat, T. L. Chenevert, C. R. Meyer, P. E. Mckeever, D. E. Hall, B. A. Hoff, T. D. Johnson, A. Rehemtulla, and B. D. Ross, "The functional diffusion map: an imaging biomarker for the early prediction of cancer treatment outcome", Neoplasia **8**, 259–267 (2006) (cit. on p. 170).

[231]D. A. Hamstra, K. C. Lee, B. A. Moffat, T. L. Chenevert, A. Rehemtulla, and B. D. Ross, "Diffusion magnetic resonance imaging: an imaging treatment response biomarker to chemoradiotherapy in a mouse model of squamous cell cancer of the head and neck", Translational Oncology **1**, 187–194 (2008) (cit. on p. 170).

[232]F. De Keyzer, V. Vandecaveye, H. Thoeny, F. Chen, Y. Ni, G. Marchal, R. Hermans, S. Nuyts, W. Landuyt, and H. Bosmans, "Dynamic contrast-enhanced and diffusion-weighted MRI for early detection of tumoral changes in single-dose and fractionated radiotherapy: evaluation in a rat rhabdomyosarcoma model", European Radiology **19**, 2663–2671 (2009) (cit. on p. 170).

[233]S. Kim, L. Loevner, H. Quon, E. Sherman, G. Weinstein, A. Kilger, and H. Poptani, "Diffusion weighted MRI for predicting and detecting early response to chemoradiation therapy of squamous cell carcinomas of the head and neck", Clinical Cancer Research **15**, 986–994 (2009) (cit. on p. 170).

[234]C. Ricci, L. Moroni, and S. Danti, "Cancer tissue engineering - new perspectives in understanding the biology of solid tumours - a critical review", OA Tissue Engineering **1**, 4 (2013) (cit. on p. 170).

[235]A. Nyga, U. Cheema, and M. Loizidou, "Tissue engineering for cancer therapy: in vitro modelling of tumour invasion.", MSc Thesis (University College London, 2011) (cit. on pp. 170, 172, 179, 204).

[236]E. Burdett, F. K. Kasper, A. G. Mikos, and J. A. Ludwig, "Engineering tumors: a tissue engineering perspective in cancer biology", Tissue Engineering. Part B, Reviews **16**, 351–359 (2010) (cit. on p. 170).

[237]T. Voskoglou-Nomikos, J. L. Pater, and L. Seymour, "Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models", Clinical Cancer Research **9**, 4227–4239 (2003) (cit. on p. 170).

[238]C. S. Szot, C. F. Buchanan, P. Gatenholm, M. N. Rylander, and J. W. Freeman, "In-vestigation of cancer cell behavior on nanofibrous scaffolds", Materials Science and Engineering: C **31**, 37–42 (2011) (cit. on p. 171).

[239]R. M. Sutherland, "Cell and environment interactions in tumor microregions: the multicell spheroid model", Science **240**, 177–184 (1988) (cit. on p. 171).

[240]A. E. Freeman and R. M. Hoffman, "In vivo-like growth of human tumors in vitro.", Proceedings of the National Academy of Sciences of the United States of America **83**, 2694–2698 (1986) (cit. on p. 171).

[241] A. Nyga, U. Cheema, and M. Loizidou, "3D tumour models: novel in vitro approaches to cancer studies", Journal of Cell Communication and Signaling **5**, 239–248 (2011) (cit. on p. 171).

[242] B. Obradovic, ed., *Cell and tissue engineering* (Springer, Berlin; Belgrade, 2011), 275 pp. (cit. on p. 171).

[243] A. Nyga, M. Loizidou, M. Emberton, and U. Cheema, "A novel tissue engineered three-dimensional in vitro colorectal cancer model", Acta Biomaterialia **9**, 7917–7926 (2013) (cit. on p. 171).

[244] H. Jaganathan, D. L. Hugar, and A. Ivanisevic, "Examining MRI contrast in three-dimensional cell culture phantoms with DNA-templated nanoparticle chains", ACS Applied Materials & Interfaces **3**, 1282–1288 (2011) (cit. on pp. 172, 204).

[245] K. P. M. Ricketts, U. Cheema, A. Nyga, A. Castoldi, C. Guazzoni, T. Magdeldin, M. Emberton, A. P. Gibson, G. J. Royle, and M. Loizidou, "A 3D in vitro cancer model as a platform for nanoparticle uptake and imaging investigations", Small **10**, 3954–3961 (2014) (cit. on p. 172).

[246] Bio-Rad Laboratories, "Comparison of count reproducibility, accuracy, and time to results between a hemocytometer and the TC10™ automated cell counter", Bulletin 6003 Rev A (2010) (cit. on p. 173).

[247] D. W. McRobbie, *MRI from picture to proton* (Cambridge University Press, 2003), 380 pp. (cit. on p. 175).

[248] M. A. Brown and R. C. Semelka, *MRI: basic principles and applications* (John Wiley & Sons, 2011), 280 pp. (cit. on p. 175).

[249] V. Kuperman, *Magnetic resonance imaging: physical principles and applications* (Academic Press, 2000), 200 pp. (cit. on p. 175).

[250] K. V. R. Chary and G. Govil, *NMR in biological systems: from molecules to human* (Springer Science & Business Media, 2008), 552 pp. (cit. on p. 176).

[251] R. A. Brown, M. Wiseman, C.-B. Chuo, U. Cheema, and S. N. Nazhat, "Ultrarapid engineering of biomimetic materials and tissues: fabrication of nano- and microstructures by plastic compression", Advanced Functional Materials **15**, 1762–1770 (2005) (cit. on p. 179).

[252] E. A. A. Neel, U. Cheema, J. C. Knowles, R. A. Brown, and S. N. Nazhat, "Use of multiple unconfined compression for control of collagen gel scaffold density and mechanical properties", Soft Matter **2**, 986–992 (2006) (cit. on p. 180).

[253] A. Purea and A. G. Webb, "Reversible and irreversible effects of chemical fixation on the NMR properties of single cells", Magnetic Resonance in Medicine **56**, 927–931 (2006) (cit. on p. 181).

[254] R. J. Dawe, D. A. Bennett, J. A. Schneider, S. K. Vasireddi, and K. Arfanakis, "Post-mortem MRI of human brain hemispheres: T2 relaxation times during formaldehyde fixation", Magnetic Resonance in Medicine **61**, 810–818 (2009) (cit. on p. 181).

[255]C. M. J. de Bazelaire, G. D. Duhamel, N. M. Rofsky, and D. C. Alsop, "MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results", Radiology **230**, 652–659 (2004) (cit. on pp. 189, 206).

[256]C. T. Rodgers, S. K. Piechnik, L. J. Delabarre, P.-F. Van de Moortele, C. J. Snyder, S. Neubauer, M. D. Robson, and J. T. Vaughan, "Inversion recovery at 7T in the human myocardium: measurement of T1, inversion efficiency and B1+", Magnetic Resonance in Medicine **70**, 1038–1046 (2013) (cit. on pp. 189, 206).

[257]S.-W. Chou, Y.-H. Shau, P.-C. Wu, Y.-S. Yang, D.-B. Shieh, and C.-C. Chen, "In vitro and in vivo studies of FePt nanoparticles for dual modal CT/MRI molecular imaging", Journal of the American Chemical Society **132**, 13270–13278 (2010) (cit. on p. 204).

[258]J. A. Brydson, *Plastics materials* (Butterworth-Heinemann, Oxford, 1999) (cit. on p. 206).

[259]R. Deichmann and A. Haase, "Quantification of T1 values by SNAPSHOT-FLASH NMR imaging", Journal of Magnetic Resonance (1969) **96**, 608–612 (1992) (cit. on p. 206).